

GenCore Version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 03:19:51 : Search time 2725.73 Seconds
(without alignments)
42.367 Million cell updates/sec

Title: US-09-380-826A-4

Perfect score: 7

Sequence: 1 tcttcca 7

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapept 1.0

Searched: 1472140 seqs, 8248589755 residues
Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Genembi: *
1: gb_ba: *
2: gb_htg: *
3: gb_in: *
4: gb_cm: *
5: gb_ov: *
6: gb_ph: *
7: gb_pat: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_cm: *
20: em_ov: *
21: em_or: *
22: em_pat: *
23: em_ph: *
24: em_pl: *
25: em_ro: *
26: em_sts: *
27: em_sy: *
28: em_un: *
29: em_vl: *
30: em_htgo_hum: *
31: em_htgo_inv: *
32: em_htgo_rnd: *
33: em_htg_hum: *
34: em_htg_inv: *
35: em_htg_rnd: *
36: em_htg_other: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	7	100.0	10	6	AR062865	AR062865 Sequence
C 2	7	100.0	12	6	AX073580	AX073580 Sequence
C 3	7	100.0	13	6	A00432	A00432 Nucleotide
C 4	7	100.0	14	6	A25812	A25812 polynucleot
C 5	7	100.0	14	6	A40518	A40518 Sequence 55
C 6	7	100.0	14	6	A88271	A88271 Sequence 41
C 7	7	100.0	14	6	A88273	A88273 Sequence 42
C 8	7	100.0	14	6	A89045	A89045 Sequence 11
C 9	7	100.0	14	6	A90238	A90238 Sequence 41
C 10	7	100.0	14	6	A90240	A90240 Sequence 42
C 11	7	100.0	14	6	AR119028	AR119028 Sequence
C 12	7	100.0	14	6	AX007733	AX007733 Sequence
C 13	7	100.0	14	6	AX007734	AX007734 Sequence
C 14	7	100.0	14	6	AX009010	AX009010 Sequence
C 15	7	100.0	14	6	I26234	I26234 Sequence 19
C 16	7	100.0	15	6	A88272	A88272 Sequence 42
C 17	7	100.0	15	6	A88466	A88466 Sequence 61
C 18	7	100.0	15	6	A90239	A90239 Sequence 42
C 19	7	100.0	15	6	A90433	A90433 Sequence 61
C 20	7	100.0	15	6	AR033435	AR033435 Sequence
C 21	7	100.0	15	6	AR033480	AR033480 Sequence
C 22	7	100.0	15	6	AR113257	AR113257 Sequence
C 23	7	100.0	15	6	AR113302	AR113302 Sequence
C 24	7	100.0	15	6	AR132941	AR132941 Sequence
C 25	7	100.0	15	6	AR132942	AR132942 Sequence
C 26	7	100.0	15	6	AR133304	AR133304 Sequence
C 27	7	100.0	15	6	AR133305	AR133305 Sequence
C 28	7	100.0	15	6	AR133306	AR133306 Sequence
C 29	7	100.0	15	6	AR133307	AR133307 Sequence
C 30	7	100.0	15	6	AR133343	AR133343 Sequence
C 31	7	100.0	15	6	AR133344	AR133344 Sequence
C 32	7	100.0	15	6	AR133345	AR133345 Sequence
C 33	7	100.0	15	6	AR133714	AR133714 Sequence
C 34	7	100.0	15	6	AX007732	AX007732 Sequence
C 35	7	100.0	15	6	AX081111	AX081111 Sequence
C 36	7	100.0	15	6	I57664	I57664 Sequence 20
C 37	7	100.0	15	6	I57709	I57709 Sequence 24
C 38	7	100.0	16	6	A96937	A96937 Sequence 15
C 39	7	100.0	16	6	A97828	A97828 Sequence 10
C 40	7	100.0	16	6	AR084439	AR084439 Sequence
C 41	7	100.0	16	6	AR093881	AR093881 Sequence
C 42	7	100.0	16	6	AX133195	AX133195 Sequence
C 43	7	100.0	16	6	I34946	I34946 Sequence 32
C 44	7	100.0	16	6	I52073	I52073 Sequence 15
C 45	7	100.0	17	6	A65762	A65762 Sequence 43

ALIGNMENTS

RESULT 1
AR062865/c 10 bp DNA PAT 29-SEP-1999
LOCUS AR062865 Sequence 13 from patent US 5843767.
DEFINITION AR062865
ACCESSION AR062865
VERSION AR062865.1 GI:5990556
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 10)
AUTHORS Beattie,K.L.

TITLE Microfabricated, flowthrough porous apparatus for discrete
detection of binding reactions
JOURNAL Patent: US 5843767-A 13 01-DEC-1998;
FEATURES Location/Qualifiers
source 1..10

BASE COUNT 4 a 4 c 0 g 2 t
ORIGIN

Query Match 100.0%; Score 7; DB 6; Length 10;
Best Local Similarity 100.0%; Pred. No. 5.1e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7
|||||
Db 7 TGTGGA 1

RESULT 2
AX073580/c 12 bp DNA PAT 06-FEB-2001
LOCUS
DEFINITION Sequence 2 from Patent WO0104320.
ACCESSION AX073580
VERSION AX073580.1 GI:12710003
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 12)
AUTHORS Schmidt,A.C., Skiadopoulos,M.H., Collins,P.L., Murphy,B.R.,
TITLE Bailey,J.E. and Durbin,A.P.
JOURNAL Attenuated human-bovine chimeric parainfluenza virus (piv) vaccines
PATENT: WO 0104320-A 2 18-JAN-2001;
THE GOVERNMENT OF THE UNITED STATES OF AMERICA (US)

FEATURES
source 1..12
Location/Qualifiers

BASE COUNT 4 a 4 c 1 g 3 t
ORIGIN
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Sequence flanking site for introduction of Sgr A1
site for BPPIV3 Ka"

Query Match 100.0%; Score 7; DB 6; Length 12;
Best Local Similarity 100.0%; Pred. No. 5e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7
|||||
Db 7 TGTGGA 1

RESULT 3
A00432/c 13 bp DNA PAT 11-FEB-1993
LOCUS
DEFINITION Nucleotide sequence 7 from patent number WO9010459.
ACCESSION A00432
VERSION A00432.1 GI:14504
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 13)
AUTHORS
TITLE PHARMACEUTICAL COMPOSITION USEFUL IN THE PREVENTION OR TREATMENT OF
JOURNAL PAPILLOMAVIRUS-INDUCED TUMOURS
PATENT: WO 9010459-A 7 20-SEP-1990;
FEATURES
source Location/Qualifiers
1..13
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 4 a 3 c 4 g 2 t
ORIGIN

Query Match 100.0%; Score 7; DB 6; Length 13;
Best Local Similarity 100.0%; Pred. No. 5e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7
|||||
Db 10 TGTGGA 4

RESULT 4
A25812/c 14 bp DNA PAT 14-MAR-1995
LOCUS
DEFINITION polynucleotide 14C12.
ACCESSION A25812
VERSION A25812.1 GI:904780
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 14)
AUTHORS
JOURNAL Patent: FR 2680520-A 19 26-FEB-1993;
FEATURES
source 1..14
Location/Qualifiers

Query Match 100.0%; Score 7; DB 6; Length 14;
Best Local Similarity 100.0%; Pred. No. 5e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7
|||||
Db 12 TGTGGA 6

RESULT 5
A40518 14 bp DNA PAT 05-MAR-1997
LOCUS
DEFINITION Sequence 55 from Patent WO9425578.
ACCESSION A40518
VERSION A40518.1 GI:2296553
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 14)
AUTHORS
TITLE ANTISENSE-OLIGONUCLEOTIDES FOR THE TREATMENT OF IMMUNOSUPPRESSIVE
JOURNAL EFFECTS OF TRANSFORMING GROWTH FACTOR--g(b) (TGF--g(b))
PATENT: WO 9425578-A 55 10-NOV-1994;
FEATURES
source Location/Qualifiers
1..14
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 3 a 3 c 4 g 4 t
ORIGIN

Query Match 100.0%; Score 7; DB 6; Length 14;
Best Local Similarity 100.0%; Pred. No. 5e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7
|||||
Db 3 TGTGGA 9

RESULT 6
A88271/c 14 bp DNA PAT 22-JAN-2000
LOCUS
DEFINITION Sequence 419 from Patent WO9833904.

ACCESSION A88271
VERSION A88271.1 GI:6736841
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 14)
AUTHORS Brysch,W. and Schlingensiepen,K.
TITLE AN ANTISENSE OLIGONUCLEOTIDE PREPARATION METHOD
JOURNAL Patent: WO 9833904-A 419 06-AUG-1998;
BIOGOSTIK GES (DE); BRYSCH WOLFGANG (DE)
FEATURES
source 1..14
/db_xref="taxon:32644"
BASE COUNT 5 a 4 c 0 g 5 t
ORIGIN

Query Match 100.0%; Score 7; DB 6; Length 14;
Best Local Similarity 100.0%; Pred. No. 5e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7
|||||
Db 11 TGTGGA 5

RESULT 7
LOCUS A88273 14 bp DNA PAT 22-JAN-2000
DEFINITION Sequence 421 from Patent WO9833904.
ACCESSION A88273
VERSION A88273.1 GI:6736843
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 14)
AUTHORS Brysch,W. and Schlingensiepen,K.
TITLE AN ANTISENSE OLIGONUCLEOTIDE PREPARATION METHOD
JOURNAL Patent: WO 9833904-A 421 06-AUG-1998;
BIOGOSTIK GES (DE); BRYSCH WOLFGANG (DE)
FEATURES
source 1..14
/db_xref="taxon:32644"
BASE COUNT 6 a 4 c 0 g 4 t
ORIGIN

Query Match 100.0%; Score 7; DB 6; Length 14;
Best Local Similarity 100.0%; Pred. No. 5e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7
|||||
Db 12 TGTGGA 6

RESULT 8
LOCUS A89045 14 bp DNA PAT 22-JAN-2000
DEFINITION Sequence 1193 from Patent WO9833904.
ACCESSION A89045
VERSION A89045.1 GI:6737615
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 14)
AUTHORS Brysch,W. and Schlingensiepen,K.
TITLE AN ANTISENSE OLIGONUCLEOTIDE PREPARATION METHOD

JOURNAL Patent: WO 9833904-A 1193 06-AUG-1998;
BIOGOSTIK GES (DE); BRYSCH WOLFGANG (DE)
FEATURES
source 1..14
/db_xref="taxon:32644"
BASE COUNT 3 a 3 c 4 g 4 t
ORIGIN

Query Match 100.0%; Score 7; DB 6; Length 14;
Best Local Similarity 100.0%; Pred. No. 5e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7
|||||
Db 3 TGTGGA 9

RESULT 9
LOCUS A90238 14 bp DNA PAT 22-JAN-2000
DEFINITION Sequence 419 from Patent EP0856579.
ACCESSION A90238
VERSION A90238.1 GI:6738752
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 14)
AUTHORS Brysch,W.D. and Schlingensiepen,K.D.
TITLE An antisense oligonucleotide preparation method
JOURNAL Patent: EP 0856579-A 419 05-AUG-1998;
BIOGOSTIK GES (DE)
FEATURES
source 1..14
/db_xref="taxon:32644"
BASE COUNT 5 a 4 c 0 g 5 t
ORIGIN

Query Match 100.0%; Score 7; DB 6; Length 14;
Best Local Similarity 100.0%; Pred. No. 5e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7
|||||
Db 11 TGTGGA 5

RESULT 10
LOCUS A90240 14 bp DNA PAT 22-JAN-2000
DEFINITION Sequence 421 from Patent EP0856579.
ACCESSION A90240
VERSION A90240.1 GI:6738754
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 14)
AUTHORS Brysch,W.D. and Schlingensiepen,K.D.
TITLE An antisense oligonucleotide preparation method
JOURNAL Patent: EP 0856579-A 421 05-AUG-1998;
BIOGOSTIK GES (DE)
FEATURES
source 1..14
/db_xref="taxon:32644"
BASE COUNT 6 a 4 c 0 g 4 t
ORIGIN

Query Match 100.0%; Score 7; DB 6; Length 14;
Best Local Similarity 100.0%; Pred. No. 5e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tgttga 7
|||||
Db 12 TGTGGA 6

RESULT 11

AR119028 AR119028 14 bp DNA PAT 16-MAY-2001
LOCUS Sequence 154 from patent US 6150092.
DEFINITION AR119028
ACCESSION AR119028
VERSION AR119028.1 GI:14100938
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 14)
AUTHORS Uchida,K., Uchida,T., Tanaka,Y., Matsuda,Y. and Kondo,S.
TITLE Antisense nucleic acid compound targeted to VEGF
JOURNAL Patent: US 6150092-A 154 21-NOV-2000;
FEATURES Location/Qualifiers
source 1..14
/organism="unknown"

BASE COUNT 2 a 1 c 6 g 5 t
ORIGIN

Query Match 100.0%; Score 7; DB 6; Length 14;
Best Local Similarity 100.0%; Pred. No. 5e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tgttga 7
|||||
Db 7 TGTGGA 13

RESULT 12
AX007733 14 bp DNA PAT 06-SEP-2000
LOCUS AX007733/c
DEFINITION Sequence 275 from Patent WO9967428.
ACCESSION AX007733
VERSION AX007733.1 GI:9995430
KEYWORDS
SOURCE
ORGANISM Aids-associated retrovirus.
Aids-associated retrovirus
VIRUSES: Retroid viruses; Retroviridae.
1 (bases 1 to 14)
AUTHORS Stuyver,L.
TITLE Method for detection of drug-selected mutations in the hiv protease gene
JOURNAL Patent: WO 9967428-A 275 29-DEC-1999;
INNOGENETICS NV (BE); STUYVER LIEVEN (BE)
FEATURES Location/Qualifiers
source 1..14
/organism="Aids-associated retrovirus"
/db_xref="taxon:11966"

BASE COUNT 5 a 6 c 0 g 3 t
ORIGIN

Query Match 100.0%; Score 7; DB 6; Length 14;
Best Local Similarity 100.0%; Pred. No. 5e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tgttga 7
|||||
Db 12 TGTGGA 6

RESULT 13
AX007734 14 bp DNA PAT 06-SEP-2000
LOCUS AX007734/c
DEFINITION Sequence 276 from Patent WO9967428.
ACCESSION AX007734
VERSION AX007734.1 GI:9995431
KEYWORDS
SOURCE
ORGANISM Aids-associated retrovirus.
Aids-associated retrovirus
VIRUSES: Retroid viruses; Retroviridae.
1 (bases 1 to 14)
AUTHORS Stuyver,L.
TITLE Method for detection of drug-selected mutations in the hiv protease gene
JOURNAL Patent: WO 9967428-A 276 29-DEC-1999;
INNOGENETICS NV (BE); STUYVER LIEVEN (BE)
FEATURES Location/Qualifiers
source 1..14
/organism="Aids-associated retrovirus"
/db_xref="taxon:11966"

BASE COUNT 5 a 6 c 0 g 3 t
ORIGIN

Query Match 100.0%; Score 7; DB 6; Length 14;
Best Local Similarity 100.0%; Pred. No. 5e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tgttga 7
|||||
Db 13 TGTGGA 7

RESULT 14
AX009010 14 bp DNA PAT 06-SEP-2000
LOCUS AX009010
DEFINITION Sequence 43 from Patent WO9963975.
ACCESSION AX009010
VERSION AX009010.1 GI:9996384
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 14)
AUTHORS Brysch,W., Schlingensiepen,K.H. and Schlingensiepen,R.
TITLE A method for stimulating the immune system
JOURNAL Patent: WO 9963975-A 43 16-DEC-1999;
BIOGENOSTIK GES (DE); BRYSCH WOLFGANG (DE); SCHLINGENSTIEPEN KARL HERMANN (DE); SCHLINGENSTIEPEN REIMAR (DE)
FEATURES Location/Qualifiers
source 1..14
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 3 a 3 c 4 g 4 t
ORIGIN

Query Match 100.0%; Score 7; DB 6; Length 14;
Best Local Similarity 100.0%; Pred. No. 5e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tgttga 7
|||||
Db 3 TGTGGA 9

RESULT 15
I26234 14 bp DNA PAT 07-OCT-1996
LOCUS I26234/c
DEFINITION Sequence 19 from patent US 5556955.
ACCESSION I26234
VERSION I26234.1 GI:1606104

KEYWORDS

SOURCE

Unknown.

ORGANISM

Unknown.

REFERENCE

1 (bases 1 to 14)

AUTHORS

Vergnaud,G.

TITLE

Process for detection of new polymorphic loci in a DNA sequence, nucleotide sequences forming hybridization probes and their applications

JOURNAL

Patent: US 5556955-A 19 17-SEP-1996;

FEATURES

Location/Qualifiers

BASE COUNT

4 a 4 c 2 g 4 t

ORIGIN

Query Match

100.0%; Score 7; DB 6; Length 14;

Best Local

Similarity 100.0%; Pred. No. 5e+05;

Matches

7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY

1 tgttga 7

111111

Db

12 TGTGGA 6

Search completed: December 15, 2001, 03:19:53

Job time: 7683 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 03:29:43 ; Search time 401.91 Seconds
(Without alignments)
14.932 Million cell updates/sec

Title: US-09-380-826a-4
Perfect score: 7
Sequence: 1 tgttga 7

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N.Geneseq_1101.*
1: /SIDS2/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7	100.0	7	19	AAV58898
2	7	100.0	10	20	AAV59938
3	7	100.0	10	21	AAZ78251
4	7	100.0	10	21	AAZ80950
5	7	100.0	10	21	AAZ82660
6	7	100.0	10	21	AAZ84297
7	7	100.0	10	22	AAH18992
8	7	100.0	10	22	AAV35297
9	7	100.0	10	22	AAZ42210
10	7	100.0	10	22	AAZ43406
11	7	100.0	10	20	AAZ7658

C 12	7	100.0	12	22	AAZ84988	BPIV3 nucleotide f
C 13	7	100.0	14	14	AAO40606	Hypervariable regi
C 14	7	100.0	14	15	AAO78406	Antisense oligonuc
C 15	7	100.0	14	19	AAV97198	Potato citrate syn
C 16	7	100.0	14	19	AAV48830	ErbB-2 gene antise
C 17	7	100.0	14	19	AAV48832	ErbB-2 gene antise
C 18	7	100.0	14	20	AAA21661	Integrin alpha 6 s
C 19	7	100.0	14	20	AAV92044	Human C-raf target
C 20	7	100.0	14	21	AAZ97785	HIV-1 protease gen
C 21	7	100.0	14	21	AAZ97785	HIV-1 protease gen
C 22	7	100.0	14	21	AAZ65483	Immunosuppressant
C 23	7	100.0	14	21	AAZ64788	Substrate for hair
C 24	7	100.0	15	2	AAAI0044	Sequence of initia
C 25	7	100.0	15	2	AAAI0045	Sequence of initia
C 26	7	100.0	15	17	AAZ66645	Human CD40 hamme
C 27	7	100.0	15	17	AAZ66275	Mouse B7-2 hamme
C 28	7	100.0	15	17	AAZ66276	Mouse B7-2 hamme
C 29	7	100.0	15	17	AAZ66274	Mouse B7-2 hamme
C 30	7	100.0	15	17	AAZ66235	Mouse B7-2 hamme
C 31	7	100.0	15	17	AAZ66236	Mouse B7-2 hamme
C 32	7	100.0	15	17	AAZ66237	Mouse B7-2 hamme
C 33	7	100.0	15	17	AAZ66238	Mouse B7-2 hamme
C 34	7	100.0	15	17	AAZ65872	Human B7-2 hamme
C 35	7	100.0	15	17	AAZ65873	Human B7-2 hamme
C 36	7	100.0	15	17	AAZ50248	Rabbit CERP HH rib
C 37	7	100.0	15	17	AAZ50250	Rabbit CERP HH rib
C 38	7	100.0	15	19	AAV49025	rb gene antisense
C 39	7	100.0	15	19	AAV48831	ErbB-2 gene antise
C 40	7	100.0	15	20	AAZ31673	Tag sequence of a
C 41	7	100.0	15	21	AAZ59285	Human NR8 gene pro
C 42	7	100.0	15	21	AAZ59298	Human NR8 gene pro
C 43	7	100.0	15	21	AAZ59299	Human NR8 gene pro
C 44	7	100.0	15	21	AAZ90823	Human NR8 gene pro
C 45	7	100.0	15	21	AAZ90839	Human NR8 gene pro

ALIGNMENTS

RESULT 1	
ID AAV58898	standard; DNA: 7 BP.
XX	
AC AAV58898;	
XX	
DT 20-JAN-1999	(first entry)
XX	
DE Leptospiira nucleotide sequence.	
XX	
KW Infection; pathogenic Leptospiira; protective immunity; therapy;	
KM diagnosis; ss.	
XX	
OS Leptospiira sp.	
XX	
PN WO9840099-A1.	
XX	
PD 17-SEP-1998.	
XX	
PF 06-MAR-1998;	98WO-AU00145.
XX	
PR 07-MAR-1997;	97AU-0005494.
XX	
PA (AGRI-) AGRIC VICTORIA SERVICES PTY LTD.	
PA (PIGR-) PIG RES & DEV CORP.	
XX	
PI Chappel RJ;	
XX	
DR WPI: 1998-520791/44.	
XX	
PT New isolated pathogenic Leptospiira bacterium - useful for, e.g	
PT developing products for conferring protective immunity, and for	
PT prophylactic or therapeutic treatment	
XX	

PS Claim 15; Page 71; 94pp; English.

XX This sequence represents a Leptospira DNA sequence isolated from the
CC pathogenic Leptospira (US) bacterium of the invention. The bacterium
CC belongs to serogroup Hurstbridge or serovar hurstbridge or the species
CC L. falnei. The US bacterium can be used for conferring protective
CC immunity against pathogenic LS bacteria in humans or animals. The
CC bacteria can also be used for prophylactic or therapeutic treatment of LS
CC infections. The DNAs and antibodies may also be used for detection and
CC diagnosis of past or present LS infection.

XX
SQ Sequence 7 BP; 1 A; 0 C; 3 G; 3 T; 0 other;

QY Query Match 100.0%; Score 7; DB 19; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.2e+08;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 tgttga 7
1 tgttga 7
Db 1 tgttga 7

RESULT 2
AAx9938/C
ID AAx9938 standard; DNA; 10 BP.

XX AAx9938;
AC
XX 21-OCT-1999 (first entry)
DT
XX Human parkin gene Intron 7 fragment.
DE
XX
XX Parkinson's disease related gene; parkin gene; variant; gene therapy;
KM Intron; ss.

XX Homo sapiens.
OS
XX WO9940191-A1.
PN
XX 12-AUG-1999.
PD
XX 09-FEB-1999; 99WO-JP00545.
PF
XX 09-FEB-1998; 98JP-0027531.
PR
XX (MIZU/) MIZUNO Y.
PA (SHIM/) SHIMIZU N.
XX
XX Mizuno Y, Shimizu N;
PI
XX WPI: 1999-494295/41.
DR
XX
XX Gene implicated in the pathology of Parkinson's disease, used for
PT treatment of the disease
PT
XX Claim 11; Page 100; 114pp; English.

XX This sequence represents a fragment of an intron from the gene of the
CC invention. The gene has been designated the parkin gene, and variants of
CC it are implicated in the pathology of Parkinson's disease, and found in
CC parkinson's disease patients. The sequences may be used for the
CC diagnosis, treatment (including gene therapy) and investigation of
CC parkinson's disease.

XX
SQ Sequence 10 BP; 3 A; 3 C; 1 G; 3 T; 0 other;

QY Query Match 100.0%; Score 7; DB 20; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.3e+04;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 tgttga 7

Db 9 TGTTCGA 3
|||||||

RESULT 3
AAz78251
ID AAz78251 standard; DNA; 10 BP.

XX AAz78251;
AC
XX 10-APR-2000 (first entry)
DT
XX Human dendritic cell SAGE tag, SEQ ID NO:679.
DE
XX
XX SAGE tag: serial analysis of gene expression; antigen-presenting cell;
KM APC; monocyte-derived dendritic cell; differential gene expression;
KM immunostimulatory cofactor; costimulatory factor; CTL;
KM cytotoxic T-lymphocyte; tumour antigen; immunotherapy; anticancer; ss.

XX Homo sapiens.
OS
XX WO9965924-A2.
PN
XX 23-DEC-1999.
PD
XX
XX 18-JUN-1999; 99WO-US13800.
PF
XX
XX 19-JUN-1998; 98US-0089833.
PR 19-JUN-1998; 98US-0089844.
PR 19-JUN-1998; 98US-0089853.
PR 19-JUN-1998; 98US-0089878.
PR 19-JUN-1998; 98US-0089991.
PR 19-JUN-1998; 98US-0089992.
PR 19-JUN-1998; 98US-0089993.
PR 19-JUN-1998; 98US-0089994.
PR 19-JUN-1998; 98US-0089997.
PR 19-JUN-1998; 98US-0089999.
PR 19-JUN-1998; 98US-0090000.
PR 19-JUN-1998; 98US-0090035.
PR 19-JUN-1998; 98US-0090036.
PR 19-JUN-1998; 98US-0090039.
PR 19-JUN-1998; 98US-0090040.
PR 19-JUN-1998; 98US-0090041.
PR 19-JUN-1998; 98US-0090042.
PR 19-JUN-1998; 98US-0090043.
PR 19-JUN-1998; 98US-0090044.
PR 19-JUN-1998; 98US-0090045.
PR 19-JUN-1998; 98US-0090047.
PR 19-JUN-1998; 98US-0090048.
PR 19-JUN-1998; 98US-0090072.
PR 19-JUN-1998; 98US-0090076.
PR 19-JUN-1998; 98US-0090077.
PR 19-JUN-1998; 98US-0090078.
PR 19-JUN-1998; 98US-0090079.
PR 19-JUN-1998; 98US-0090080.
PR 08-DEC-1998; 98US-0111715.

XX (GENZ) GENZYME CORP.
PA (ROBE/) ROBERTS B L.
PA (SHAN/) SHANKARA S.
XX
XX Roberts BL, Shankara S;
PI
XX WPI: 2000-106077/09.
DR
XX
XX Isolated polynucleotides differentially expressed in antigen-presenting
PT cells, useful in gene vaccines against cancer -
PT
XX Claim 1; Page 84; 130pp; English.

XX Sequences AA277573-279709 represent SAGE (serial analysis of gene
CC expression) tags used to identify mRNA transcripts encoding
CC immunostimulatory cofactor proteins which are preferentially or

CC differentially expressed in monocyte-derived dendritic cells compared
CC with monocytes. Some of the transcripts correspond to known genes or
CC ESTs (expressed sequence tags) which were previously unknown to be
CC preferentially or differentially expressed in dendritic cells, while
CC other transcripts correspond to novel genes. Antigen-presenting cell
CC (APC)-associated costimulatory factors play an important role in the
CC activation of the cytotoxic immune response, particularly against tumour
CC cells. Tumour antigen presentation via the MHC (major histocompatibility
CC complex) and subsequent recognition by T-cell receptors is alone
CC insufficient to activate a robust cytotoxic immune response that can
CC lyse the tumour cells, immunostimulatory cofactors also being required
CC for efficient activation of cytotoxic T-lymphocytes (CTLs). Nucleic acid
CC sequences identified using the SAGE tags have several potential uses.
CC They may be used in vaccines to induce an immune response, particularly
CC against a tumour antigen; to modulate the genotype of an APC; to screen
CC for agents that modulate expression of differentially expressed genes in
CC an APC; and as hybridisation probes/amplification primers for the
CC diagnosis, prognosis and monitoring of diseases related to abnormal
CC expression of these genes. Detection of the dendritic cell
CC differentially expressed genes, or of their encoded proteins, can be used
CC to identify cells as belonging to the monocyte lineage. Cells containing
CC these genes can be used in active immunotherapy (or to stimulate
CC production of a population of antigen-specific effector cells) and
CC vectors containing them are used in gene therapy. Co-administration of
CC tumour antigens and APC-associated costimulatory factors ensures adequate
CC antigen presentation to endogenous APCs and upregulates the APCs for the
CC presentation of co-stimulatory signals, migration to T cell-rich sites,
CC secretion of T cell growth factors and secretion of chemokines for
CC recruitment of immune effector cells.

Query Match	100.0%	Score 7:	DB 21:	length 10:
Best Local Similarity	100.0%	Pred. No.	3.3e+04:	
Matches 7; Conservative	0;	Mismatches	0;	Indels 0;
				Gaps 0;

	Qy	Db
1	tgttga	7
2		8
3	tgttga	9

RESULT	4
AA280950	
ID	AA280950 standard; DNA; 10 BP.

AC AAZ80950;

DT 07-APR-2000 (first entry)

DE Metastatic breast tumour cell upregulated transcript tag #184.

KM Human; metastatic breast tumour tissue; breast cancer; tag; primer
KM non-metastatic breast tumour tissue; gene therapy; anticancer;
KM antimetastatic; vaccine; diagnosis; ss.

OS Homo sapiens.

PN W09965928-A2.

PD 23-DEC-1999

PF 18-JUN-1999; 99WO-US13647.

PR 19-JUN-1998; 98US-0089853.

PR 19-JUN-1998; 98US-00900339.

PR 19-JUN-1998; 98US-0090041.

PA (GENZ) GENZYME CORP

PA (SHAN/) SHANKARA S.

XX Roberts BL, Shankara S;
PI
XX WPI; 2000-106079/09.
DR

PT Isolated polynucleotides differentially expressed between metastatic
PT and non-metastatic breast cancer cells, useful for diagnosis,
PT prevention and treatment of cancer -

PS Claim 1; Page 63; 219pp; English.

AA28076 to AA283941 represent tags corresponding to distinct transcripts that are preferentially transcribed in the metastatic breast tumour tissue (i.e. are upregulated in metastatic breast tumour cells). AA283942 to AA286677 represent tags corresponding to distinct transcripts that are preferentially transcribed in the primary or non-metastatic breast tumour tissue (i.e. are downregulated in metastatic breast tumour cells). These transcripts can be used for diagnosis, prognosis, monitoring and treatment of breast cancer, particularly where metastatic diagnosis is by standard immunoassays or hybridisation/amplification reactions. Compounds that modulate expression of the transcripts are potentially useful for treatment of (metastatic) breast cancer, while promoters from the transcripts are used to direct expression, in selected cell types, of e.g. therapeutic genes (also ribozymes or antisense sequences), particularly an antigen-encoding sequence for use in gene or cell-based vaccines. Polypeptides encoded by the transcripts are also useful in vaccines, for diagnosing breast cancer and for raising specific antibodies (Ab). Ab are used to detect the polypeptides or as therapeutic agents. Host cells that produce the polypeptides can be used to expand and isolate populations of educated, antigen-specific immune effector cells, e.g. cytotoxic T lymphocytes, and these used for adoptive immunotherapy.

Sequence 10 BP; 2 A; 0 C; 4 G; 4 T; 0 other;

Query Match	100.0%	Score 7	DB 21	Length 10
Best Local Similarity	100.0%	Pred. No.	3.3e+04	
Matches 7	Conservative 0	Mismatches 0	Indels 0	Gaps 0

Qy	1	tgttga	7
Db	4	tgttga	10

RESULT	5
AAZ82660	
ID	AAZ82660 standard; DNA; 10 BP

AC AAZ82660;

DT 07-APR-2000 (first entry)

DE Metastatic breast tumour cell upregulated transcript tag #1894.

KW Human; metastatic breast tumour tissue; breast cancer; tag; primer; non-metastatic breast tissue; gene therapy; anticancer

KW antimetastatic; vaccine; diagnosis; ss.
 XY

OS Homo sapiens.
XY

PN W09965928-A2.
XX

PD 23-DEC-1999
XX

PE 18-JUN-1999; 99WO-USL3647
XX

PR 19-JUN-1998; 9805-0089853.
PR 19-JUN-1998; 9805-0089853.
PR 19-JUN-1998; 9805-0089853.

PR	19-JUN-1998;	98US-0090039.
PR	19-JUN-1998;	98US-0090040.

PR 19-JUN-1998; 9805-0090041:
XX

PA (GEN2) GENZYME CORP.
 PA (ROBE/) ROBERTS B L.
 PA (SHAN/) SHANKARA S.
 XX
 XX
 PI Roberts BL, Shankara S;
 XX
 XX WPI: 2000-106079/09.
 DR
 XX
 PT Isolated polynucleotides differentially expressed between metastatic
 PT and non-metastatic breast cancer cells, useful for diagnosis,
 PT prevention and treatment of cancer -
 XX
 XX Claim 1: Page 110; 219pp; English.
 PS
 XX
 CC AA280767 to AA283941 represent tags corresponding to distinct
 CC transcripts that are preferentially transcribed in the metastatic breast
 CC tumour tissue (i.e. are upregulated in metastatic breast tumour cells).
 CC AA283942 to AA286677 represent tags corresponding to distinct transcripts
 CC that are preferentially transcribed in the primary or non-metastatic
 CC breast tumour tissue (i.e. are downregulated in metastatic breast tumour
 CC cells). These transcripts can be used for diagnosis, prognosis,
 CC monitoring and treatment of breast cancer, particularly where metastatic.
 CC diagnosis is by standard immunoassays or hybridisation/amplification
 CC reactions. Compounds that modulate expression of the transcripts are
 CC potentially useful for treatment of (metastatic) breast cancer, while
 CC promoters from the transcripts are used to direct expression, in selected
 CC cell types, of e.g. therapeutic genes (also ribozymes or antisense
 CC sequences), particularly an antigen-encoding sequence for use in gene or
 CC cell-based vaccines. Polypeptides encoded by the transcripts are also
 CC useful in vaccines; for diagnosing breast cancer and for raising
 CC specific antibodies (Ab). Ab are used to detect the polypeptides or as
 CC therapeutic agents. Host cells that produce the polypeptides can be used
 CC to expand and isolate populations of educated, antigen-specific immune
 CC effector cells, e.g. cytotoxic T lymphocytes, and these used for
 CC adoptive immunotherapy.
 CC
 XX Sequence 10 BP; 2 A; 0 C; 4 G; 4 T; 0 other;
 SQ

Query Match 100.0%; Score 7; DB 21; Length 10;
 Best Local Similarity 100.0%; Pred. No. 3.3e+04;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tgttga 7
 |||||
 Db 2 tgttga 8

RESULT 6
 AA284297
 ID AA284297 standard; DNA; 10 BP.
 XX
 AC AA284297;
 XX
 XX 07-APR-2000 (first entry)
 DT
 XX
 DE Metastatic breast tumour cell downregulated transcript tag #3531.
 XX
 XX Human: metastatic breast tumour tissue: breast cancer; tag: primer;
 KW non-metastatic breast tumour tissue; gene therapy; anticancer;
 KW antimetastatic; vaccine; diagnosis; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO965928-A2.
 XX
 PD 23-DEC-1999.
 XX
 PF 18-JUN-1999; 99WO-US13647.
 XX
 PR 19-JUN-1998; 98US-0089853.
 PR 19-JUN-1998; 98US-0089997.
 PR 19-JUN-1998; 98US-0090039.

PR 19-JUN-1998; 98US-0090040.
 PR 19-JUN-1998; 98US-0090041.
 XX
 XX
 PA (GEN2) GENZYME CORP.
 PA (ROBE/) ROBERTS B L.
 PA (SHAN/) SHANKARA S.
 XX
 XX
 PI Roberts BL, Shankara S;
 XX
 XX WPI: 2000-106079/09.
 DR
 XX
 PT Isolated polynucleotides differentially expressed between metastatic
 PT and non-metastatic breast cancer cells, useful for diagnosis,
 PT prevention and treatment of cancer -
 XX
 XX Claim 1: Page 153; 219pp; English.
 PS
 XX
 CC AA280767 to AA283941 represent tags corresponding to distinct
 CC transcripts that are preferentially transcribed in the metastatic breast
 CC tumour tissue (i.e. are upregulated in metastatic breast tumour cells).
 CC AA283942 to AA286677 represent tags corresponding to distinct transcripts
 CC that are preferentially transcribed in the primary or non-metastatic
 CC breast tumour tissue (i.e. are downregulated in metastatic breast tumour
 CC cells). These transcripts can be used for diagnosis, prognosis,
 CC monitoring and treatment of breast cancer, particularly where metastatic.
 CC diagnosis is by standard immunoassays or hybridisation/amplification
 CC reactions. Compounds that modulate expression of the transcripts are
 CC potentially useful for treatment of (metastatic) breast cancer, while
 CC promoters from the transcripts are used to direct expression, in selected
 CC cell types, of e.g. therapeutic genes (also ribozymes or antisense
 CC sequences), particularly an antigen-encoding sequence for use in gene or
 CC cell-based vaccines. Polypeptides encoded by the transcripts are also
 CC useful in vaccines; for diagnosing breast cancer and for raising
 CC specific antibodies (Ab). Ab are used to detect the polypeptides or as
 CC therapeutic agents. Host cells that produce the polypeptides can be used
 CC to expand and isolate populations of educated, antigen-specific immune
 CC effector cells, e.g. cytotoxic T lymphocytes, and these used for
 CC adoptive immunotherapy.
 CC
 XX Sequence 10 BP; 2 A; 0 C; 3 G; 5 T; 0 other;
 SQ

Query Match 100.0%; Score 7; DB 21; Length 10;
 Best Local Similarity 100.0%; Pred. No. 3.3e+04;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tgttga 7
 |||||
 Db 2 tgttga 8

RESULT 7
 AAH18992
 ID AAH18992 standard; DNA; 10 BP.
 XX
 AC AAH18992;
 XX
 XX 21-JUN-2001 (first entry)
 DT
 XX
 DE UCP3 polymorphism detection allele specific primer #105.
 XX
 XX UCP3; uncoupling protein 3; polymorphism; obesity;
 KW diabetes mellitus; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200118232-A2.
 XX
 PD 15-MAR-2001.
 XX
 PF 08-SEP-2000; 2000WO-US24784.
 PF 08-SEP-1999; 99US-0152789.
 PR

XX (GENA-) GENAISSANCE PHARM INC.
PA (STEP/) STEPHENS J C.
XX
PI Chew A, Choi JY, Denton RR, Nandabalan K;
XX WPI: 2001-218562/22.
DR
XX Nucleic acids encoding uncoupling protein 3 (mitochondrial, proton
PT carrier) (UCP3) proteins comprising single nucleotide polymorphisms,
PT useful for the design of drugs for treating obesity -
XX
PS Disclosure: Page 23: 94pp: English.
XX
CC The present invention relates to the human uncoupling protein 3
CC (mitochondrial, proton carrier) (UCP3) gene and polymorphisms.
CC The polymorphisms are associated with obesity, especially
CC diabetes mellitus associated obesity. They polymorphisms may be
CC identified and analysed to determine whether an individual is
CC susceptible to obesity and may be used as the basis for targeted
CC design of drugs to treat obesity. The present sequence was used in
CC the identification and amplification of UCP3 polymorphisms.
XX
SQ Sequence 10 BP; 2 A; 1 C; 3 G; 4 T; 0 other;

Query Match 100.0%; Score 7; DB 22; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.3e+04;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tgttga 7
Db 1 tgttga 7

RESULT 8
AAF35297/c
ID AAF35297 standard; DNA; 10 BP.
XX
AC AAF35297;
XX
DT 23-MAR-2001 (first entry)
XX
DE Yeast NORF gene SAGE tag oligonucleotide SEQ ID NO:2036.
XX
KW Yeast; Saccharomyces cerevisiae; characterisation: cell cycle: NORF;
KW not previously assigned open reading frame; nonannotated ORF; SAGE;
KW serial analysis of gene expression; antifungal; tag; Identification;
KW linker; PCR primer; ds.
XX
OS Saccharomyces cerevisiae.
XX
PN WO200077214-A2.
XX
PD 21-DEC-2000.
XX
PF 14-JUN-2000; 2000WO-US16223.
XX
PR 16-JUN-1999; 99US-0335032.
XX
PA (UJO) UNIV JOHNS HOPKINS.
XX
PI Velculescu V, Vogelstein B, Kinzler K;
XX WPI: 2001-061874/07.
XX
DR
XX Yeast gene coding sequences comprising NORF genes with serial analysis
PT of gene expression (SAGE) tags, useful for studying, monitoring and
PT affecting phases of the cell cycle -
XX
PS Example: Page 72: 419pp: English.
XX The present invention describes an isolated DNA molecule comprising a

CC coding sequence of a yeast gene selected from a group of 745 NORF (not
CC previously assigned open reading frame; or nonannotated ORF) genes
CC comprising a SAGE (serial analysis of gene expression) tag. Also
CC described are: (1) a method (M1) of using NORF genes to affect the cell
CC cycle comprising administering a NORF gene whose expression varies by at
CC least 10% between any two phases of the cell cycle selected from log
CC phase, S phase and G2/M; (2) a method (M2) for screening candidate
CC antifungal drugs comprising: (a) contacting a test substance with a
CC yeast cell; and (b) monitoring expression of a NORF gene whose
CC expression varies as in M1, where a test substance which modifies the
CC expression of the yeast gene is a candidate antifungal drug; (3) a method
CC (M3) for identifying human genes which are involved in cell cycle
CC progression comprising contacting human DNA with a probe which comprises
CC at least 10 contiguous nucleotides of a NORF gene whose expression varies
CC as in M1; and (4) a method (M4) for identifying a candidate drug as a
CC member of a class of drugs having a characteristic effect on gene
CC expression in a yeast cell comprising contacting a yeast cell with a
CC candidate drug and monitoring expression in the yeast cell of at least 1
CC NORF gene whose expression is affected by the class of drugs. The NORF
CC genes may be used to study, monitor and affect phases of the cell cycle,
CC the differentially expressed genes may be used as markers of phases of
CC the cell cycle. The methods may be used to identify candidate drugs which
CC affect the cell cycle and for identification of antifungal drugs.
CC AAF33268 to AAF44064 represent SAGE tags used in the exemplification of
CC the present invention. AAF33262 to AAF33267 represent linkers and PCR
CC primers used in the SAGE method, in the exemplification of the present
XX invention.
XX
SQ Sequence 10 BP; 4 A; 4 C; 0 G; 2 T; 0 other;

Query Match 100.0%; Score 7; DB 22; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.3e+04;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tgttga 7
Db 10 TGTTCGA 4

RESULT 9
AAF42210/c
ID AAF42210 standard; DNA; 10 BP.
XX
AC AAF42210;
XX
DT 23-MAR-2001 (first entry)
XX
DE Yeast NORF gene SAGE tag oligonucleotide SEQ ID NO:8949.
XX
KW Yeast; Saccharomyces cerevisiae; characterisation: cell cycle: NORF;
KW not previously assigned open reading frame; nonannotated ORF; SAGE;
KW serial analysis of gene expression; antifungal; tag; Identification;
KW linker; PCR primer; ds.
XX
OS Saccharomyces cerevisiae.
XX
PN WO200077214-A2.
XX
PD 21-DEC-2000.
XX
PF 14-JUN-2000; 2000WO-US16223.
XX
PR 16-JUN-1999; 99US-0335032.
XX
PA (UJO) UNIV JOHNS HOPKINS.
XX
PI Velculescu V, Vogelstein B, Kinzler K;
XX WPI: 2001-061874/07.
XX
DR
XX Yeast gene coding sequences comprising NORF genes with serial analysis
PT of gene expression (SAGE) tags, useful for studying, monitoring and

PT affecting phases of the cell cycle -
XX
XX Example; Page 319; 419pp; English.
XX
CC The present invention describes an isolated DNA molecule comprising a
CC coding sequence of a yeast gene selected from a group of 745 NORF (not
CC previously assigned open reading frame; or nonannotated ORF) genes
CC comprising a SAGE (serial analysis of gene expression) tag. Also
CC described are: (1) a method (M1) of using NORF genes to affect the cell
CC cycle comprising administering a NORF gene whose expression varies by at
CC least 10% between any two phases of the cell cycle selected from 10g
CC phase, S phase and G2/M; (2) a method (M2) for screening candidate
CC antifungal drugs comprising: (a) contacting a test substance with a
CC yeast cell; and (b) monitoring expression of a NORF gene whose
CC expression varies as in M1, where a test substance which modifies the
CC expression of the yeast gene is a candidate antifungal drug; (3) a method
CC (M3) for identifying human genes which are involved in cell cycle
CC progression comprising contacting human DNA with a probe which comprises
CC at least 10 contiguous nucleotides of a NORF gene whose expression varies
CC as in M1; and (4) a method (M4) for identifying a candidate drug as a
CC member of a class of drugs having a characteristic effect on gene
CC expression in a yeast cell comprising contacting a yeast cell with a
CC candidate drug and monitoring expression in the yeast cell of at least 1
CC NORF gene whose expression is affected by the class of drugs. The NORF
CC genes may be used to study, monitor and affect phases of the cell cycle,
CC the differentially expressed genes may be used as markers of phases of
CC the cell cycle. The methods may be used to identify candidate drugs which
CC affect the cell cycle and for identification of antifungal drugs.
CC AAF33268 to AAF44064 represent SAGE tags used in the exemplification of
CC the present invention. AAF33262 to AAF33267 represent linkers and PCR
CC primers used in the SAGE method, in the exemplification of the present
CC invention.
XX
SQ Sequence 10 BP; 4 A; 3 C; 1 G; 2 T; 0 other;
Query Match 100.0%; Score 7; DB 22; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.3e+04;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 tgttga 7
|||||
Db 10 TGTTCGA 4
RESULT 10
AAF43406
ID AAF43406 standard; DNA; 10. BP.
XX
AC AAF43406;
XX
XX 23-MAR-2001 (first entry)
XX
DE Yeast NORF gene SAGE tag oligonucleotide SEQ ID NO:11545.
XX
XX Yeast; Saccharomyces cerevisiae; characterisation; cell cycle; NORF;
KW nor previously assigned open reading frame; nonannotated ORF; SAGE;
KW serial analysis of gene expression; antifungal; tag; identification;
KW linker; PCR primer; ds.
XX
XX Saccharomyces cerevisiae.
OS
XX WO200077214-A2.
PN
XX 21-DEC-2000.
XX
XX 14-JUN-2000; 2000WO-US16223.4
PF
XX 16-JUN-1999; 99US-0335032.
PR
XX (UYZO) UNTV JOHNS HOPKINS.
PA
XX Velculescu V, Vogelstein B, Kinzler K;
PI

XX
DR WPI: 2001-061874/07.
XX
PT Yeast gene coding sequences comprising NORF genes with serial analysis
PT of gene expression (SAGE) tags, useful for studying, monitoring and
PT affecting phases of the cell cycle -
XX
XX Example; Page 362; 419pp; English.
XX
CC The present invention describes an isolated DNA molecule comprising a
CC coding sequence of a yeast gene selected from a group of 745 NORF (not
CC previously assigned open reading frame; or nonannotated ORF) genes
CC comprising a SAGE (serial analysis of gene expression) tag. Also
CC described are: (1) a method (M1) of using NORF genes to affect the cell
CC cycle comprising administering a NORF gene whose expression varies by at
CC least 10% between any two phases of the cell cycle selected from 10g
CC phase, S phase and G2/M; (2) a method (M2) for screening candidate
CC antifungal drugs comprising: (a) contacting a test substance with a
CC yeast cell; and (b) monitoring expression of a NORF gene whose
CC expression varies as in M1, where a test substance which modifies the
CC expression of the yeast gene is a candidate antifungal drug; (3) a method
CC (M3) for identifying human genes which are involved in cell cycle
CC progression comprising contacting human DNA with a probe which comprises
CC at least 10 contiguous nucleotides of a NORF gene whose expression varies
CC as in M1; and (4) a method (M4) for identifying a candidate drug as a
CC member of a class of drugs having a characteristic effect on gene
CC expression in a yeast cell comprising contacting a yeast cell with a
CC candidate drug and monitoring expression in the yeast cell of at least 1
CC NORF gene whose expression is affected by the class of drugs. The NORF
CC genes may be used to study, monitor and affect phases of the cell cycle,
CC the differentially expressed genes may be used as markers of phases of
CC the cell cycle. The methods may be used to identify candidate drugs which
CC affect the cell cycle and for identification of antifungal drugs.
CC AAF33268 to AAF44064 represent SAGE tags used in the exemplification of
CC the present invention. AAF33262 to AAF33267 represent linkers and PCR
CC primers used in the SAGE method, in the exemplification of the present
CC invention.
XX
SQ Sequence 10 BP; 1 A; 0 C; 6 G; 3 T; 0 other;
Query Match 100.0%; Score 7; DB 22; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.3e+04;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 tgttga 7
|||||
Db 1 tgttga 7
RESULT 11
AAx77658/C
ID AAx77658 standard; DNA; 11 BP.
XX
AC AAx77658;
XX
XX 09-AUG-1999 (first entry)
DT
XX
DE N11 active EGS 22.
XX
XX External guide sequence; EGS; target mRNA; identification; diagnostic;
KW inactivation; essential gene; therapy; ss.
XX
XX Synthetic.
OS
XX WO9927135-A2.
PN
XX 03-JUN-1999.
XX
XX 20-NOV-1998; 98WO-US24854.
PF
XX 30-MAR-1998; 98US-0079851.
PR
XX 21-NOV-1997; 97US-0976220.
PR

Mon Dec 17 07:48:26 2001

us-09-380-826a-4.rng

Page 7

```

XX (INNO-) INNOVIR LAB INC.
XX
XX
PI Kindt TJ, Nilsen TW, Robertson HD,
XX
XX WPI: 1999-357853/30.
XX
PT Identifying and inhibiting functional nucleic acid molecules in
PT cells
XX
XX Example 3; Page 28; 58pp; English.
XX
XX This invention describes a novel method allowing essential or functional
XX genes to be rapidly identified and inactivated. The method is able to
XX firstly identify most of the essential genes in an organism (i.e. a
XX bacteria or a eukaryote) needed for survival, and secondly it provides
XX for reducing or inactivating their expression. The method is able to
XX identify functional oligonucleotide molecules able to be used as
XX diagnostic reagents and therapeutics. The method provides a means for
XX identifying essential genes whose sequence is known only as part of a
XX genome with unknown function, as well as a means for identifying
XX functional oligonucleotide molecules. The method involves the use of a
XX nucleic acid molecule comprising (a) a first reporter gene encoding a
XX fusion protein comprising a protein of interest (itself translated from
XX an RNA of interest) and a reporter protein, a second reporter gene
XX encoding a second reporter protein, and (c) a targeting guide sequence
XX functional oligonucleotide molecule such as an external guide sequence
XX (EGS), a ribozyme or an antisense RNA and targeted to the RNA of interest
XX at a site on the first reporter gene able to encode the RNA of interest.
XX
SQ Sequence 11 BP; 4 A; 4 C; 1 G; 2 T; 0 other;

Query Match      100.0%; Score 7; DB 20; Length 11;
Best Local Similarity 100.0%; Pred. No. 3.3e+04;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttcga 7
   |||||||
DB 10 TGTTCGA 4

RESULT 12
AAC84988/c
ID AAC84988 standard; DNA; 12 BP.
XX
XX AAC84988;
AC
XX
XX 08-MAY-2001 (first entry)
DT
XX
XX BPIV3 nucleotide fragment.
DE
XX
XX Parainfluenza virus; PIV; infectious; human; nucleocapsid protein; BPIV;
KW nucleocapsid phosphoprotein; polymerase protein; human; bovine; HPIV;
KW chimeric; vaccine; immune response; HPIV1; HPIV2; HPIV3; BPIV3; ds.
XX
XX Bovine parainfluenza virus.
OS
XX
XX MO200104320-A1.
PN
XX
XX 18-JAN-2001.
PD
XX
XX 16-JUN-2000; 2000WO-US17066.
PF
XX
XX 09-JUL-1999; 99US-0143134.
PR
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA
XX
XX Schmidt AC, Skiadopoulos MH, Collins PL, Murphy BR, Bailly JE;
PI Durbin AP;
XX
XX WPI: 2001-081053/09.
XX

```

```

PT Isolated human-bovine chimeric parainfluenza virus (PIV), useful in an
PT attenuated vaccine to elicits an immune response against one or more
PT virus(es) selected from human PIV1 (HPIV1), HPIV2 and HPIV3 -
XX
XX
XX Example 4; Page 80; 148pp; English.
XX
XX The invention relates to an isolated human-bovine chimeric parainfluenza
XX virus (PIV) that is infectious and attenuated in humans. The virus
XX comprises a major nucleocapsid protein, a nucleocapsid phosphoprotein,
XX a large polymerase protein, and a partial or complete PIV background
XX genome, or antigenome of a human PIV (HPIV) or bovine PIV (BPIV),
XX combined with one or more heterologous gene(s) or genome segment(s) of a
XX different PIV to form a human-bovine chimeric PIV genome or antigenome.
XX The chimeric PIV is useful in attenuated vaccines to elicit immune
XX response against one or more virus(es) selected from HPIV1, HPIV2 and
XX HPIV3 and most preferably against HPIV3. The present sequence represents
XX a BPIV3 parent nucleotide fragment. This sequence is changed in a
XX recombinant BPIV3 for constructing a chimeric PIV.
XX
SQ Sequence 12 BP; 4 A; 4 C; 1 G; 3 T; 0 other;

Query Match      100.0%; Score 7; DB 22; Length 12;
Best Local Similarity 100.0%; Pred. No. 3.3e+04;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttcga 7
   |||||||
DB 7 TGTTCGA 1

RESULT 13
AAQ40606/c
ID AAQ40606 standard; DNA; 14 BP.
XX
XX AAQ40606;
AC
XX
XX 10-AUG-1993 (first entry)
DT
XX
XX Hypervariable region detection probe 14C12.
DE
XX
XX HVR; human; animal; forensic science; paternity testing; diagnosis;
KW animal breeding; hereditary diseases; tumours; allele; loss;
KW chromosomal regions; tumour region identification; ss.
XX
XX Synthetic.
OS
XX
XX FR2680520-A.
PN
XX
XX 26-FEB-1993.
PD
XX
XX 22-AUG-1991; 91FR-0010516.
PF
XX
XX 22-AUG-1991; 91FR-0010516.
PR
XX
XX (ETFR ) EPAT FRANCAIS.
PA
XX
XX Vergnaud G;
PI
XX
XX WPI: 1993-136548/17.
DR
XX
XX Detecting the hypervariable regions of DNA for diagnosing
PT hereditary illnesses and tumours - by hybridising labelled
PT polynucleotides and analysing genomic DNA of individuals which
PT react with restriction fragments
XX
XX Example; Page 13; 46pp; French.
XX
XX The sequence is that of a polynucleotide probe which may be used in
XX the detection of new hypervariable regions (HVR) in a DNA sequence.
XX HVR represent a fingerprint useful in e.g. forensic science.
XX paternity testing, animal breeding, etc. The probe may be used as
XX part of a method for the efficient detection in humans or other
XX

```

CC animals, without the use of mini-satellites or primary enrichment.
XX
SQ Sequence 14 BP; 4 A; 4 C; 2 G; 4 T; 0 other;

Query Match 100.0%; Score 7; DB 14; Length 14;
Best Local Similarity 100.0%; Pred. No. 3.3e+04;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7
| | | | |
DB 12 TGTGGA 6

RESULT 14
AAQ78406
ID AAQ78406 standard; DNA; 14 BP.
XX
AC AAQ78406;
XX
DT 27-JUN-1995 (first entry)
XX
DE Antisense oligonucleotide hybridising to TGF-beta gene.

XX Transforming growth factor beta; TGF-beta; antisense; treatment;
KM tumour; angiogenesis; breast tumour; neurofibroma; glioma;
KM glioblastoma; carcinogenesis; carcinoma; oesophagus; oesophageal;
KM gastric; gut; immunosuppression; oligonucleotide; ss.

XX Synthetic.

OS

XX

PN MO9425588-A.

PD 10-NOV-1994.

XX

PF 29-APR-1994; 94MO-EP01362.

XX

PR 30-APR-1993; 93EP-0107089.

PR 13-MAY-1993; 93EP-0107849.

XX

PA (BIOG-) BIOGNOSTIK GES BIOMOLEKULARE DIAGNOSTIK.

XX

PI Bogdahn U, Brysch W, Schlingensiepen G, Schlingensiepen K;

PI Schlingensiepen R;

XX

DR WPI; 1994-358266/44.

XX

PT New transforming growth factor beta anti:sense

PT oligo:nucleotide(s) - for treating immunosuppression, tumours,

XX etc.

PS Claim 6; Page 40; 74pp; English.

XX

CC The antisense oligonucleotides are useful in the treatment of

CC tumours in which expression of TGF-beta is of relevance for

CC pathogenetically and/or inhibition of pathological angiogenesis. They

CC are used especially for the treatment of the immunosuppressive

CC effect of TGF-beta, augmentation of the proliferation of cytotoxic

CC lymphocytes, treatment of endogenous hyperexpression of TGF-beta,

CC treatment of breast tumours, neurofibromas and malignant gliomas,

CC including glioblastomas, treatment and prophylaxis of skin

CC carcinogenesis, and treatment of oesophageal and gastric carcinomas.

CC See AAQ78352-078488. The sequences given in GENESQ files

CC AAQ78352-078407 and AAQ78408 are antisense oligodeoxynucleotides of

CC TGF-beta 1. The sequences given in GENESQ files AAQ78408-78487 are

CC antisense oligodeoxynucleotides of TGF-beta 2 in the form of

CC phosphorothioate analogues.

CC
XX
SQ Sequence 14 BP; 3 A; 3 C; 4 G; 4 T; 0 other;

Query Match 100.0%; Score 7; DB 15; Length 14;
Best Local Similarity 100.0%; Pred. No. 3.3e+04;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7
| | | | |
DB 3 tgttga 9

RESULT 15
AAV97198/C
ID AAV97198 standard; RNA; 14 BP.
XX
AC AAV97198;
XX
DT 01-MAR-1999 (first entry)
XX
DE Potato citrate synthase target sequence position 123.

XX Solanidine; glucosyltransferase; potato; citrate synthase; target;
KM hammerhead ribozyme; hairpin ribozyme; alkaloid biosynthesis;
KM flower formation; cleavage; solanaceous plant; ss.

XX Solanum tuberosum.

XX

PN WO9832843-A2.

PD 30-JUL-1998.

XX

PF 14-JAN-1998; 98MO-US00738.

XX

PR 24-NOV-1997; 97US-0979416.

PR 28-JAN-1997; 97US-0036545.

XX

PA (RIBO-) RIBOZYME PHARM INC.

XX

PI McSwiggen JA, Zwick MG;

PI WPI; 1998-427939/36.

DR

XX

PT New enzymatic nucleic acid(s) - useful for, e.g. reducing alkaloid

PT biosynthesis or regulating flowering

XX

PS Claim 54; Page 59; 79pp; English.

XX

CC The present invention describes enzymatic nucleic acid molecules with

CC RNA-cleaving activity (e.g. ribozymes) which are capable of modulating

CC the expression of plant genes: (i) involved in biosynthesis of

CC alkaloids; or (ii) involved in flower formation. AAV95982 to AAV96334,

CC and AAV96335 to AAV96354 represent potato solanidine glucosyltransferase

CC hammerhead and hairpin ribozymes, respectively. AAV95981,

CC and AAV96355 to AAV96734 represent potato solanidine glucosyltransferase

CC target sequences. AAV96773 to AAV97170, and AAV97171 to AAV97195

CC represent potato citrate synthase hammerhead and hairpin ribozymes,
CC respectively. AAV96735 to AAV96772, and AAV97196 to AAV97220 represent
CC potato citrate synthase target sequences. Ribozymes of the present
CC invention can be used to inhibit the synthesis of toxic alkaloids in
CC solanaceous plants, particularly potato but also tomato, pepper,
CC aubergine and datura or to inhibit flowering in potato, lettuce, spinach,
CC cabbage, brussel sprouts, arugula, kale, collards, chard, beet, turnip,
CC sweet potato and turf grass. Also the ribozymes can be used for RNA
CC manipulation in the same way that restriction endonucleases are for DNA,
CC as well as to examine genetic drift and mutations in plants and to
CC detect specific RNA. The ribozymes can be targeted to specific genes or
CC to consensus sequences within a family of related genes, and being
CC catalytic need to be present at only very low concentrations.

CC
XX
SQ Sequence 14 BP; 4 A; 4 C; 5 G; 1 U; 0 other;

Query Match 100.0%; Score 7; DB 19; Length 14;
Best Local Similarity 100.0%; Pred. No. 3.3e+04;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Mon Dec 17 07:48:26 2001

us-09-380-826a-4.rng

Page 9

OY 1 t9lt9ga 7
| | | | |
Db 13 TGTTCGA 7

Search completed: December 15, 2001, 03:29:44
Job time: 7884 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 03:22:46 : Search time 172.39 Seconds
(without alignments)
9.196 Million cell updates/sec

Title: US-09-380-826a-4

Perfect score: 7

Sequence: 1 tgttga 7

Scoring table: IDENTITY_NUC

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCBUS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	7	100.0	10	US-08-631-751A-13	Sequence 13, Appl
C 2	7	100.0	11	US-09-136-523-25	Sequence 25, Appl
C 3	7	100.0	14	US-08-303-004-19	Sequence 19, Appl
C 4	7	100.0	14	US-08-765-340-154	Sequence 154, Appl
C 5	7	100.0	15	US-08-182-668A-201	Sequence 201, Appl
C 6	7	100.0	15	US-08-182-668A-246	Sequence 246, Appl
C 7	7	100.0	15	US-08-363-240A-614	Sequence 614, Appl
C 8	7	100.0	15	US-08-363-240A-615	Sequence 615, Appl
C 9	7	100.0	15	US-08-363-240A-616	Sequence 616, Appl
C 10	7	100.0	15	US-08-774-306A-201	Sequence 201, Appl
C 11	7	100.0	15	US-08-774-306A-246	Sequence 246, Appl
C 12	7	100.0	15	US-08-585-684B-1366	Sequence 1366, Appl
C 13	7	100.0	15	US-08-585-684B-1367	Sequence 1367, Appl
C 14	7	100.0	15	US-08-585-684B-1729	Sequence 1729, Appl
C 15	7	100.0	15	US-08-585-684B-1730	Sequence 1730, Appl
C 16	7	100.0	15	US-08-585-684B-1731	Sequence 1731, Appl
C 17	7	100.0	15	US-08-585-684B-1732	Sequence 1732, Appl
C 18	7	100.0	15	US-08-585-684B-1768	Sequence 1768, Appl
C 19	7	100.0	15	US-08-585-684B-1769	Sequence 1769, Appl
C 20	7	100.0	15	US-08-585-684B-1770	Sequence 1770, Appl
C 21	7	100.0	15	US-08-585-684B-2139	Sequence 2139, Appl
C 22	7	100.0	15	US-09-064-156A-201	Sequence 201, Appl
C 23	7	100.0	15	US-09-064-156A-246	Sequence 246, Appl
C 24	7	100.0	15	US-09-038-073-1366	Sequence 1366, Appl
C 25	7	100.0	15	US-09-038-073-1367	Sequence 1367, Appl
C 26	7	100.0	15	US-09-038-073-1729	Sequence 1729, Appl
C 27	7	100.0	15	US-09-038-073-1730	Sequence 1730, Appl

C 28	7	100.0	15	US-09-038-073-1731	Sequence 1731, Appl
C 29	7	100.0	15	US-09-038-073-1732	Sequence 1732, Appl
C 30	7	100.0	15	US-09-038-073-1768	Sequence 1768, Appl
C 31	7	100.0	15	US-09-038-073-1769	Sequence 1769, Appl
C 32	7	100.0	15	US-09-038-073-1770	Sequence 1770, Appl
C 33	7	100.0	15	US-09-038-073-2139	Sequence 2139, Appl
C 34	7	100.0	16	US-08-435-350-433	Sequence 433, Appl
C 35	7	100.0	16	US-08-166-664-15	Sequence 15, Appl
C 36	7	100.0	16	US-08-469-461-15	Sequence 15, Appl
C 37	7	100.0	16	US-07-890-609-15	Sequence 15, Appl
C 38	7	100.0	17	US-08-072-063-4	Sequence 4, Appl
C 39	7	100.0	17	US-08-119-773-23	Sequence 23, Appl
C 40	7	100.0	17	US-08-281-940-25	Sequence 25, Appl
C 41	7	100.0	17	US-08-390-850-433	Sequence 433, Appl
C 42	7	100.0	17	US-08-064-693-4	Sequence 4, Appl
C 43	7	100.0	17	US-08-373-124A-810	Sequence 810, Appl
C 44	7	100.0	17	US-08-441-370-2	Sequence 2, Appl
C 45	7	100.0	17	US-09-282-146-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-08-631-751A-13/C
Sequence 13, Application US/08631751A
Patent No. 5843767
APPLICANT: Beattie, Kenneth L.
TITLE OF INVENTION: Microfabricated, Flowthrough Porous Apparatus for Discrete Detection of Blinding Reactions
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Vinson & Elkins
STREET: 1455 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20004-1008
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/631,751A
FILING DATE: 11-Apr-11-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sanzo, Michael A.
REGISTRATION NUMBER: 36,912
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)639-6500
TELEFAX: (202)639-6604
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
US-08-631-751A-13

Query Match 100.0%; Score 7; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 6.7e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 tgttga 7
|||||
Db 7 tgttga 1

RESULT 2
US-09-196-523-25/c
Sequence 25, Application US/09196523A
Patent No. 6248525
GENERAL INFORMATION:
APPLICANT: Nilsen, Timothy W.
TITLE OF INVENTION: Method for Identifying and Inactivating Essential or
FILE REFERENCE: ILI 130
CURRENT APPLICATION NUMBER: US/09/196,523A
CURRENT FILING DATE: 1998-11-20
EARLIER APPLICATION NUMBER: 60/079,851
EARLIER FILING DATE: 1998-03-30
NUMBER OF SEQ ID NOS: 32
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 25
LENGTH: 11
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:
US-09-196-523-25

Query Match 100.0%; Score 7; DB 4; Length 11;
Best Local Similarity 100.0%; Pred. No. 6.7e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tgttga 7
|||||
DB 10 TGTGGA 4

RESULT 3
US-08-303-004-19/c
Sequence 19, Application US/08303004
Patent No. 5556955
GENERAL INFORMATION:
APPLICANT: Vergnaud, Gilles
TITLE OF INVENTION: Process for Detection of New Polymor-
TITLE OF INVENTION: phic LocI in an ADN Sequence, Nucleotide Sequences Forming
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Oliff & Herridge
STREET: P.O. Box 19928
CITY: Alexandria
STATE: Virginia
COUNTRY: U.S.A
ZIP: 22320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/303,004
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/931,311B
FILING DATE: 19920818
ATTORNEY/AGENT INFORMATION:
NAME: Herridge, William P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: WPB 28264
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6400
TELEFAX: (703) 836-2787
TELEX: 90-1799 PTO ALEX
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:

LENGTH: 14 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-303-004-19

Query Match 100.0%; Score 7; DB 1; Length 14;
Best Local Similarity 100.0%; Pred. No. 6.8e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tgttga 7
|||||
DB 12 TGTGGA 6

RESULT 4
US-08-765-340-154
Sequence 154, Application US/08765340
Patent No. 6150092
GENERAL INFORMATION:
APPLICANT: UCHIDA, K.
APPLICANT: UCHIDA, T.
APPLICANT: TANAKA, Y.
APPLICANT: MATSUDA, Y.
APPLICANT: KONDO, S.
TITLE OF INVENTION: AN ANTISENSE NUCLEIC ACID
NUMBER OF SEQUENCES: 185
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/765,340
FILING DATE: 23-DEC-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 145146/94
FILING DATE: 27-JUN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 31130/94
FILING DATE: 21-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: SERUNIAN, LESLIE
REGISTRATION NUMBER: 35,353
REFERENCE/DOCKET NUMBER: 1452-4005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
INFORMATION FOR SEQ ID NO: 154:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "synthetic DNA"
US-08-765-340-154

Query Match 100.0%; Score 7; DB 3; Length 14;

Best Local Similarity 100.0%; Pred. No. 6.8e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 ttttga 7
Db 7 ttttga 13

RESULT 5

US-08-182-968A-201
Sequence 201, Application US/08182968A
Patent No. 5610054
GENERAL INFORMATION:
APPLICANT: Draper, Kenneth G.
TITLE OF INVENTION: METHOD AND REAGENT FOR
TITLE OF INVENTION: INHIBITING HEPATITIS C
TITLE OF INVENTION: VIRUS REPLICATION
NUMBER OF SEQUENCES: 497
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/182,968A
FILING DATE: 13-JANUARY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/882,888
FILING DATE: 14-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 205/277
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 201:
SEQUENCE CHARACTERISTICS:
LENGTH: 15
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-182-968A-201

Query Match 100.0%; Score 7; DB 1; Length 15;
Best Local Similarity 57.1%; Pred. No. 6.8e+03;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ttttga 7
Db 5 ucugga 11

RESULT 6

US-08-182-968A-246/C
Sequence 246, Application US/08182968A
Patent No. 5610054
GENERAL INFORMATION:
APPLICANT: Draper, Kenneth G.
TITLE OF INVENTION: METHOD AND REAGENT FOR
TITLE OF INVENTION: INHIBITING HEPATITIS C
TITLE OF INVENTION: VIRUS REPLICATION

NUMBER OF SEQUENCES: 497
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/182,968A
FILING DATE: 13-JANUARY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/882,888
FILING DATE: 14-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 205/277
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 246:
SEQUENCE CHARACTERISTICS:
LENGTH: 15
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-182-968A-246

Query Match 100.0%; Score 7; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 6.8e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ttttga 7
Db 13 ttttga 7

RESULT 7

US-08-363-240A-614/C
Sequence 614, Application US/08363240A
Patent No. 5705388
GENERAL INFORMATION:
APPLICANT: Couture, Larry
APPLICANT: McSwigen, James
APPLICANT: Bisgaier, Charles
APPLICANT: Pape, Michael
TITLE OF INVENTION: METHOD AND REAGENT FOR
TITLE OF INVENTION: PREVENTION, INHIBITION OF
TITLE OF INVENTION: PROGRESSION AND REGRESSION
TITLE OF INVENTION: OF VASCULAR DISEASES
NUMBER OF SEQUENCES: 1243
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible

OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/363,240A
FILING DATE: December 23, 1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 210/096
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ. ID NO: 614:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-363-240A-614

Query Match 100.0%; Score 7; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 6.8e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7
|111111
Db 15 TGTGGA 9

RESULT 8
US-08-363-240A-615/C
Sequence 615, Application US/08363240A
Patent No. 5705388
GENERAL INFORMATION:
APPLICANT: Couture, Larry
APPLICANT: McSwigen, James
APPLICANT: Bisgaler, Charles
APPLICANT: Pepe, Michael
TITLE OF INVENTION: METHOD AND REAGENT FOR
TITLE OF INVENTION: PREVENTION, INHIBITION OF
TITLE OF INVENTION: PROGRESSION AND REGRESSION
TITLE OF INVENTION: OF VASCULAR DISEASES
NUMBER OF SEQUENCES: 1243
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/363,240A
FILING DATE: December 23, 1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 210/096
TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ. ID NO: 615:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-363-240A-615

Query Match 100.0%; Score 7; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 6.8e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7
|111111
Db 15 TGTGGA 9

RESULT 9
US-08-363-240A-616/C
Sequence 616, Application US/08363240A
Patent No. 5705388
GENERAL INFORMATION:
APPLICANT: Couture, Larry
APPLICANT: McSwigen, James
APPLICANT: Bisgaler, Charles
APPLICANT: Pepe, Michael
TITLE OF INVENTION: METHOD AND REAGENT FOR
TITLE OF INVENTION: PREVENTION, INHIBITION OF
TITLE OF INVENTION: PROGRESSION AND REGRESSION
TITLE OF INVENTION: OF VASCULAR DISEASES
NUMBER OF SEQUENCES: 1243
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/363,240A
FILING DATE: December 23, 1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 210/096
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ. ID NO: 616:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-363-240A-616

Query Match 100.0%; Score 7; DB 1; Length 15;

Best Local Similarity 100.0%; Pred. No. 6.8e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 tqtgga 7
111111
Db 13 ttttga 7

RESULT 10

US-08-774-306A-201
Sequence 201, Application US/08774306A
Patent No. 5869253
GENERAL INFORMATION:
APPLICANT: Draper, Kenneth G.
TITLE OF INVENTION: METHOD AND REAGENT FOR
TITLE OF INVENTION: INHIBITING HEPATITIS C
TITLE OF INVENTION: VIRUS REPLICATION
NUMBER OF SEQUENCES: 497
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Suite 4700
STATE: Los Angeles
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/774,306A
FILING DATE: December 26, 1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/182,968
FILING DATE: January 13, 1994
APPLICATION NUMBER: 07/882,888
FILING DATE: May 14, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 223/227
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 201:
SEQUENCE CHARACTERISTICS:
LENGTH: 15
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-774-306A-201

Query Match 100.0%; Score 7; DB 2; Length 15;
Best Local Similarity 57.1%; Pred. No. 6.8e+03;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 1 tqtgga 7
111111
Db 5 ucucga 11

RESULT 11

US-08-774-306A-246/C
Sequence 246, Application US/08774306A
Patent No. 5869253
GENERAL INFORMATION:
APPLICANT: Draper, Kenneth G.
TITLE OF INVENTION: METHOD AND REAGENT FOR

TITLE OF INVENTION: INHIBITING HEPATITIS C
TITLE OF INVENTION: VIRUS REPLICATION
NUMBER OF SEQUENCES: 497
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Suite 4700
STATE: Los Angeles
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/774,306A
FILING DATE: December 26, 1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/182,968
FILING DATE: January 13, 1994
APPLICATION NUMBER: 07/882,888
FILING DATE: May 14, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 223/227
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 246:
SEQUENCE CHARACTERISTICS:
LENGTH: 15
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-774-306A-246

Query Match 100.0%; Score 7; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 6.8e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 tqtgga 7
111111
Db 13 ttttga 7

RESULT 12

US-08-585-684B-1366/C
Sequence 1366, Application US/08585684B
Patent No. 5877021
GENERAL INFORMATION:
APPLICANT: Stinchcomb, Daniel T.
APPLICANT: Jarvis, Thale
TITLE OF INVENTION: MCSwigen, James
TITLE OF INVENTION: METHOD AND REAGENT FOR THE
TITLE OF INVENTION: INDUCTION OF GRAFT TOLERANCE
TITLE OF INVENTION: AND REVERSAL OF IMMUNE RESPONSES
NUMBER OF SEQUENCES: 2751
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Suite 4700
STATE: Los Angeles
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/585,684B
FILING DATE: January 16, 1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/000,951
FILING DATE: July 7, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 218/078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 1366:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-585-684B-1366

Query Match 100.0%; Score 7; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 6.8e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 tgltaga 7
|||||
Db 14 TGTTCGA 8

RESULT 13
US-08-585-684B-1367/c
Sequence 1367, Application US/08585684B
Patent No. 5877021
GENERAL INFORMATION:
APPLICANT: Stinchcomb, Daniel T.
APPLICANT: Jarvis, Thale
APPLICANT: McSwigen, James
TITLE OF INVENTION: METHOD AND REAGENT FOR THE
TITLE OF INVENTION: INDUCTION OF GRAFT TOLERANCE
TITLE OF INVENTION: AND REVERSAL OF IMMUNE RESPONSES
NUMBER OF SEQUENCES: 2751
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/585,684B
FILING DATE: January 16, 1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/000,951
FILING DATE: July 7, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 218/078
TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 1367:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-585-684B-1367

Query Match 100.0%; Score 7; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 6.8e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 tgltaga 7
|||||
Db 13 TGTTCGA 7

RESULT 14
US-08-585-684B-1729/c
Sequence 1729, Application US/08585684B
Patent No. 5877021
GENERAL INFORMATION:
APPLICANT: Stinchcomb, Daniel T.
APPLICANT: Jarvis, Thale
APPLICANT: McSwigen, James
TITLE OF INVENTION: METHOD AND REAGENT FOR THE
TITLE OF INVENTION: INDUCTION OF GRAFT TOLERANCE
TITLE OF INVENTION: AND REVERSAL OF IMMUNE RESPONSES
NUMBER OF SEQUENCES: 2751
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/585,684B
FILING DATE: January 16, 1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/000,951
FILING DATE: July 7, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 218/078
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 1729:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-585-684B-1729

Query Match 100.0%; Score 7; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 6.8e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7
 1111111
 Db 13 TGTGGA 7

RESULT 15
 US-08-585-684B-1730/c
 : Sequence 1730, Application US/08585684B
 : Patent No. 5877021
 : GENERAL INFORMATION:
 : APPLICANT: Stinchcomb, Daniel T.
 : APPLICANT: Jarvis, Thale
 : APPLICANT: McSwiggen, James
 : TITLE OF INVENTION: METHOD AND REAGENT FOR THE
 : TITLE OF INVENTION: INDUCTION OF CRAFT TOLERANCE
 : TITLE OF INVENTION: AND REVERSAL OF IMMUNE RESPONSES
 : NUMBER OF SEQUENCES: 2751
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Lyon & Lyon
 : STREET: 633 West Fifth Street
 : STREET: Suite 4700
 : CITY: Los Angeles
 : STATE: California
 : COUNTRY: U.S.A.
 : ZIP: 90071
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 : MEDIUM TYPE: storage
 : COMPUTER: IBM Compatible
 : OPERATING SYSTEM: IBM P.C. DOS 5.0
 : SOFTWARE: FastSeq Version 1.5
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/585,684B
 : FILING DATE: January 16, 1996
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 60/000,951
 : FILING DATE: July 7, 1995
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Warburg, Richard
 : REGISTRATION NUMBER: 32,327
 : REFERENCE/DOCKET NUMBER: 218/078
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (213) 489-1600
 : TELEFAX: (213) 955-0440
 : TELEX: 67-3510
 : INFORMATION FOR SEQ ID NO: 1730:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 15 base pairs
 : TYPE: nucleic acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 : US-08-585-684B-1730

Query Match 100.0%; Score 7; DB 2; Length 15;
 Best Local Similarity 100.0%; Pred. No. 6.8e+03;
 Matches 7; Conservative 0; Mismatches 0; Indels 0;
 Gaps 0;
 OY 1 tgttga 7
 1111111
 Db 13 TGTGGA 7

Search completed: December 15, 2001, 03:22:47
 Job time: 7562 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 02:33:37 : Search time 4441.54 Seconds
(without alignments)
16.936 Million cell updates/sec

Title: US-09-380-826a-4

Perfect score: 1 tglgtga 7

Sequence: 1 tglgtga 7

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues
Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST: *
1: em_estfun: *
2: em_esthum: *
3: em_estln: *
4: em_estom: *
5: em_estpl: *
6: em_estba: *
7: em_estro: *
8: em_estov: *
9: em_hlc: *
10: gb_est1: *
11: gb_est2: *
12: gb_hlc: *
13: gb_gss: *
14: em_gss_fun: *
15: em_gss_hum: *
16: em_gss_inv: *
17: em_gss_pln: *
18: em_gss_pro: *
19: em_gss_rtd: *
20: em_gss_vrt: *
21: em_gss_other: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7	100.0	19	13	AZ386406 1M0145C22
2	7	100.0	19	13	AZ2994163 2M0279F05
3	7	100.0	20	13	AZ484701 1M0311C24
4	7	100.0	20	13	AZ807221 2M0068G20
5	7	100.0	20	13	BH000596 2M0288N06
6	7	100.0	21	13	AZ361153 1M0104A17
7	7	100.0	21	13	AZ450829 1M0249E13
8	7	100.0	21	13	AZ508369 1M0350A24
9	7	100.0	22	13	TA175C05P
10	7	100.0	23	13	AZ663995 1M0543M22
11	7	100.0	23	13	AZ819914 2M0091N16
12	7	100.0	24	13	AZ999558 2M0286N24

13	7	100.0	24	13	TA143H100	AL467041 T. brucei
14	7	100.0	24	13	TA206G04P	AL475439 T. brucei
15	7	100.0	25	10	AI074857	AI074857 oy27b11.s
16	7	100.0	25	10	AI158948	AI158948 u142e07.x
17	7	100.0	25	10	AI453394	AI453394 t3j7902.x
18	7	100.0	25	13	AZ869367	AZ869367 2M0181P03
19	7	100.0	26	13	AZ352012	AZ352012 1M0090M13
20	7	100.0	26	13	AZ637522	AZ637522 1M0496N23
21	7	100.0	27	13	TA346G04Q	AL496254 T. brucei
22	7	100.0	28	13	AZ777241	AZ777241 2M0011G02
23	7	100.0	29	13	AZ332711	AZ332711 1M0061N14
24	7	100.0	29	13	AZ419519	AZ419519 1M0196N03
25	7	100.0	29	13	AZ939009	AZ939009 2M0197L13
26	7	100.0	30	13	AZ580757	AZ580757 1M0369D07
27	7	100.0	30	13	AZ829102	AZ829102 2M0106K02
28	7	100.0	30	13	AZ942800	AZ942800 2M0203D12
29	7	100.0	30	13	TA160D01P	AL473241 T. brucei
30	7	100.0	31	13	TA251B07P	AL483453 T. brucei
31	7	100.0	31	13	AZ333172	AZ333172 1M0062B09
32	7	100.0	31	13	AZ357647	AZ357647 1M0099D19
33	7	100.0	31	13	TA140G08Q	AL466834 T. brucei
34	7	100.0	32	10	AV833440	AV833440 AV833440
35	7	100.0	32	13	AZ320254	AZ320254 1M0040P07
36	7	100.0	32	13	AZ328463	AZ328463 1M0052D18
37	7	100.0	32	13	AZ467843	AZ467843 1M0279F06
38	7	100.0	32	13	AZ587241	AZ587241 1M0394D14
39	7	100.0	32	13	AZ607412	AZ607412 1M0429A24
40	7	100.0	32	13	TA187G01Q	AL476327 T. brucei
41	7	100.0	33	11	H48291	H48291 y977a01.r1
42	7	100.0	34	13	AZ408169	AZ408169 1M0179C17
43	7	100.0	34	13	AZ764028	AZ764028 1M0059G13
44	7	100.0	34	13	AZ775415	AZ775415 2M0007M16
45	7	100.0	34	13	AZ820498	AZ820498 2M0092C01

ALIGNMENTS

RESULT 1
LOCUS AZ386406 19 bp DNA
DEFINITION 1M0145C22P Mouse 10kb plasmid UGCC1M library Mus musculus genomic
clone UGCC1M0145C22 F, DNA sequence.
ACCESSION AZ386406
VERSION AZ386406.1 GI:10500106
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 19)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamli,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
'M., Rose,M., Rose,R., Stokes,R., Tinney,A., von Niederhausern,A.
and Wright,D., Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLG, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0145 row: C column: 22
Seq primer: CGTGTAAACAGACGCCACGT
Class: plasmid ends
High quality sequence stop: 19.
FEATURES
source 1..19
Location/Qualifiers

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0145C22"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. coli strain XL10-gold, T1-resistant, F-"
/note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g114732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT
5 a 2 c 6 g 6 t

ORIGIN

Query Match 100.0%; Score 7; DB 13; Length 19;
Best Local Similarity 100.0%; Pred. No. 3.6e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 tgttga 7
||||||
Db 11 TGTGTGA 17

RESULT 2
A2994163 19 bp DNA GSS 27-APR-2001
LOCUS A2994163/c
DEFINITION 2M0279F05R Mouse 10kb plasmid UUGC2M library Mus musculus genomic
clone UUGC2M0279F05 R, DNA sequence.

ACCESSION A2994163.1 GI:13865390

VERSION A2994163.1

KEYWORDS GSS.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 19)

REFERENCE 1 (bases 1 to 19)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.

Mouse whole genome scaffolding with paired end reads from 10kb

TITLE plasmid inserts

JOURNAL Unpublished (2000)

COMMENT Contact: Robert B. Weiss

University of Utah

84112 USA

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0279 row: F column: 05

Seq primer: CACACGAAACAGCTATGACG

Class: plasmid ends

High quality sequence stop: 19.

Location/Qualifiers

source

1. 19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0279F05"
/clone_lib="Mouse 10kb plasmid UUGC2M library"
/sex="Female"
/lab_host="E. coli strain XL10-gold, T1-resistant, F-"
/note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (female) was obtained from the Jackson Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g114732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT
8 a 5 c 3 g 3 t

ORIGIN

Query Match 100.0%; Score 7; DB 13; Length 19;
Best Local Similarity 100.0%; Pred. No. 3.6e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 tgttga 7
||||||
Db 11 TGTGTGA 5

RESULT 3
A2484701 20 bp DNA GSS 05-OCT-2000
LOCUS A2484701
DEFINITION 1M0311C24F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0311C24 F, DNA sequence.

ACCESSION A2484701.1 GI:10649799

VERSION A2484701

KEYWORDS GSS.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 20)

REFERENCE 1 (bases 1 to 20)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.

Mouse whole genome scaffolding with paired end reads from 10kb

TITLE plasmid inserts

JOURNAL Unpublished (2000)

COMMENT Contact: Robert B. Weiss

University of Utah

84112 USA

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0311 row: C column: 24

Seq primer: CTTGTAAACGACGCGCAGT

Class: plasmid ends

High quality sequence stop: 20.

FEATURES

source

Location/Qualifiers

```
1. .20
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="U06C1M0311C24"
/clone.lib="Mouse 10kb plasmid U06C1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-."
/notes="Vector: PMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PMD42 (g114732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."
```

BASE COUNT

```
2 a 4 c 6 g 8 t
```

ORIGIN

```
Query Match 100.0%; Score 7; DB 13; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.6e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 tgttga 7
|111111
Db 8 TGTTCGA 14
```

RESULT 4

LOCUS A2807221 20 bp DNA GSS 20-FEB-2001
DEFINITION 2M0069G20R Mouse 10kb plasmid U06C1M library Mus musculus genomic
clone U06C2M0069G20 R, DNA sequence.

ACCESSION A2807221
VERSION A2807221.1 GI:12971352
KEYWORDS GSS.

SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 20)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert length: 10000 Std Error: 0.00
Plate: 0069 row: G column: 20
Seq primer: CACACAGAAACAGCTATGACC
Class: plasmid ends

FEATURES

source

High quality sequence stop: 20.
Location/Qualifiers

```
1. .20
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="U06C2M0069G20"
/clone.lib="Mouse 10kb plasmid U06C1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-."
/notes="Vector: PMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PMD42 (g114732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."
```

BASE COUNT

```
4 a 2 c 4 g 10 t
```

ORIGIN

```
Query Match 100.0%; Score 7; DB 13; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.6e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 tgttga 7
|111111
Db 13 TGTTCGA 19
```

RESULT 5

LOCUS BH000596 20 bp DNA GSS 27-APR-2001
DEFINITION 2M0286N06R Mouse 10kb plasmid U06C2M library Mus musculus genomic
clone U06C2M0286N06 R, DNA sequence.

ACCESSION BH000596
VERSION BH000596.1 GI:13871822
KEYWORDS GSS.

SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 20)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert length: 10000 Std Error: 0.00
Plate: 0288 row: N column: 06
Seq primer: CACACAGAAACAGCTATGACC

Class: plasmid ends
High quality sequence stop: 20.
Location/Qualifiers
1. 20

FEATURES
source
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUCG2M0288N06"
/clone_lib="Mouse 10kb plasmid UUCG2M library"
/sex="Female"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (female) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g1473211419b1AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT
2 a 3 c 8 g 7 t
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 7; DB 13; Length 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 t g t t g g a 7
1111111
Db 1 T G T T G C A 7

RESULT 6
LOCUS A2361153 21 bp DNA GSS 02-OCT-2000
DEFINITION IM0104A17R Mouse 10kb plasmid UUCG1M library Mus musculus genomic
ACCESSION A2361153
VERSION A2361153.1 GI:10474853
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 21)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
TITLE Unpublished (2000)
JOURNAL
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel.: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0104 row: A column: 17

Seq primer: CACACGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 21.
Location/Qualifiers
1. 21

FEATURES
source
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUCG1M0104A17"
/clone_lib="Mouse 10kb plasmid UUCG1M library"
/sex="Male"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g1473211419b1AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT
10 a 5 c 0 g 6 t
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 7; DB 13; Length 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 t g t t g g a 7
1111111
Db 13 T G T T G C A 7

RESULT 7
LOCUS A2450829/C 21 bp DNA GSS 04-OCT-2000
DEFINITION IM0249E13R Mouse 10kb plasmid UUCG1M library Mus musculus genomic
ACCESSION A2450829
VERSION A2450829.1 GI:10606020
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 21)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
TITLE Unpublished (2000)
JOURNAL
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel.: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00

Plate: 0249 row: E column: 13
Seq primer: CACACAGAAACACGCTATGACC
Class: plasmid ends
High quality sequence stop: 21.
Location/Qualifiers

FEATURES
Source

1. 21
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUCG1M0249P13"
/clone_lib="Mouse 10kb plasmid UUCG1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g114732114(gb)AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT
ORIGIN

8 a 8 c 0 g 5 t

Query Match 100.0%; Score 7; DB 13; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.6e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 tgttga 7
|||||
Db 19 TGTGGA 13

RESULT 8
LOCUS A2508369 21 bp DNA GSS 05-OCT-2000
DEFINITION IM0350A24R Mouse 10kb plasmid UUCG1M library Mus musculus genomic clone UUCG1M0350A24 R, DNA sequence.
ACCESSION A2508369
VERSION A2508369.1 GI:10689781
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muride; Murinae; Mus.
1 (bases 1 to 21)
REFERENCE 1
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLG, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00
Plate: 0350 row: A column: 24
Seq primer: CACACAGAAACACGCTATGACC
Class: plasmid ends
High quality sequence stop: 21.
Location/Qualifiers

FEATURES
Source

1. 21
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUCG1M0350A24"
/clone_lib="Mouse 10kb plasmid UUCG1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g114732114(gb)AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT
ORIGIN

5 a 1 c 7 g 8 t

Query Match 100.0%; Score 7; DB 13; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.6e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 tgttga 7
|||||
Db 3 TGTGGA 9

RESULT 9
LOCUS TAI75C05P/c 22 bp DNA GSS 13-DEC-2000
DEFINITION T. brucei sheared genomic DNA clone 175c05, forward sequence, genomic survey sequence.
ACCESSION A1475173
VERSION A1475173.1 GI:11840185
KEYWORDS GSS.
SOURCE Trypanosoma brucei.
ORGANISM Trypanosoma brucei.
Eukaryota; Euzlenozoa; Kinetoplastida; Trypanosomatidae;
Trypanosoma.
1 (bases 1 to 22)
REFERENCE 1
AUTHORS Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R., Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L., Melville,S.E., Rajandream,M.A. and Barrell,B.G.
TITLE Direct Submission
JOURNAL Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and nh@sanger.ac.uk
COMMENT Constructed at the Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from a cloned population of Trypanosoma brucei (TREN927/4 Gnat 10.1) was mechanically sheared to give a tight size distribution (4 kb). The v + i method used for the library construction is described in detail in Smith, H. and Venter, J.C. (Making small

Insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Barrell, Oxford University Press, 1999).

Email: nelsayed@cligr.org
Details of T. brucei sequencing at the Sanger Centre are available at <http://www.sanger.ac.uk/projects/T-brucei/>.

Location/Qualifiers

1. 22
/organism="Trypanosoma brucei"

/strain="TREU927"

/db_xref="taxon:5691"

/clone="175c05"

BASE COUNT 4 a 6 c 3 g 9 t

ORIGIN

Query Match 100.0%; Score 7; DB 13; Length 22;
Best Local Similarity 100.0%; Pred. No. 3.6e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tcttga 7
| | | | |
Db 13 TCTTGA 7

RESULT 10

AZ663995 23 bp DNA GSS 14-DEC-2000

LOCUS 1M0543M22R Mouse 10kb plasmid UUGC1M library Mus musculus genomic

DEFINITION clone UUGC1M0543M22 R, DNA sequence.

ACCESSION AZ663995.1 GI:11801141

VERSION GSS.

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES
SOURCE

1. 23
/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUGC1M0543M22"

/clone.lib="Mouse 10kb plasmid UUGC1M library"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"

/note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA

was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g147321141pb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 9 a 8 c 1 g 5 t

ORIGIN

Query Match 100.0%; Score 7; DB 13; Length 23;
Best Local Similarity 100.0%; Pred. No. 3.6e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tcttga 7
| | | | |
Db 17 TCTTGA 11

RESULT 11

AZ819914 23 bp DNA GSS 20-FEB-2001

LOCUS 2M0091N16R Mouse 10kb plasmid UUGC1M library Mus musculus genomic

DEFINITION clone UUGC2M0091N16 R, DNA sequence.

ACCESSION AZ819914 GI:12989822

VERSION GSS.

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES
SOURCE

1. 23
/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUGC2M0091N16"

/clone.lib="Mouse 10kb plasmid UUGC1M library"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"

/note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g114732114|gblAF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT
ORIGIN
5 a 7 c 4 g 7 t

Query Match 100.0%; Score 7; DB 13; Length 23;
Best Local Similarity 100.0%; Pred. No. 3.6e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 tcttga 7
|||||
Db 22 TGTGGA 16

RESULT 12

A2999558 24 bp DNA GSS 27-APR-2001
LOCUS 2M0286N2AR Mouse 10kb plasmid UUGC2M library Mus musculus genomic
DEFINITION clone UUGC2M0286N24 R, DNA sequence.

ACCESSION A2999558
VERSION A2999558.1 GI:13870785
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 24) "Baider, M., Beacorn, T., Duval, B., Hamli, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tinney, A., von Niederhausern, A. and Wright, D., Weiss, R.

AUTHORS

Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
Unpublished (2000)

JOURNAL COMMENT
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA

Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0286 row: N column: 24
Seq primer: CACACGACGAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 24.

FEATURES
SOURCE

1. .24
Location/Qualifiers
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="UUGC2M0286N24"
/clone_lib="Mouse 10kb plasmid UUGC2M library"
/sex="Female"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: pMD42nv; Purified genomic DNA from M. musculus C57BL/6J (female) was obtained from the Jackson Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA

was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g114732114|gblAF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT
ORIGIN
2 a 5 c 8 g 9 t

Query Match 100.0%; Score 7; DB 13; Length 24;
Best Local Similarity 100.0%; Pred. No. 3.7e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 tcttga 7
|||||
Db 17 TGTGGA 23

RESULT 13

TA143H10Q 24 bp DNA GSS 13-DEC-2000
LOCUS T. brucei sheared genomic DNA clone 143h10, reverse sequence,
DEFINITION genomic survey sequence.

ACCESSION A1467041
VERSION A1467041.1 GI:11836396
KEYWORDS GSS.
SOURCE Trypanosoma brucei.
ORGANISM Trypanosoma brucei

REFERENCE 1 (bases 1 to 24) "Hall, N., Bowman, S., Lennard, N.J., Doggett, J., Atkin, R., Chillingworth, C., Ormond, D., Harris, B., El-Sayed, N., Hou, L., Melville, S.E., Rajandream, M.A. and Barrell, B.G.

AUTHORS

Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and nh@sanger.ac.uk

Constructed at the Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from a cloned population of Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared to give a tight size distribution (4 kb). The v + i method used for the library construction is described in detail in Smith, H. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaubin and B. Barrell, Oxford University Press, 1999).

JOURNAL COMMENT

Email: nelsayed@tigr.org
Details of T. brucei sequencing at the Sanger Centre are available at http://www.sanger.ac.uk/Projects/T_brucei/.

FEATURES
SOURCE

1. .24
Location/Qualifiers
/organism="Trypanosoma brucei"
/strain="TREU927"
/db_xref="taxon:5691"
/clone="143h10"
BASE COUNT 2 a 3 c 9 g 10 t
ORIGIN
Query Match 100.0%; Score 7; DB 13; Length 24;
Best Local Similarity 100.0%; Pred. No. 3.7e+05;

```

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 tgttga 7
Db 12 TGTGGA 18

RESULT 14
LOCUS TA206G04P/C 24 bp DNA GSS 13-DEC-2000
DEFINITION T. brucei sheared genomic DNA clone 206g04, forward sequence,
genomic survey sequence.
ACCESSION AL475439
VERSION AL475439.1 GI:11842420
KEYWORDS GSS.
SOURCE Trypanosoma brucei.
ORGANISM Trypanosoma brucei
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
Trypanosoma.
REFERENCE 1 (bases 1 to 24)
AUTHORS Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R.,
Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L.,
Melville,S.E., Rajandream,M.A. and Barrell,B.G.
DIRECT SUBMISSION
JOURNAL Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
nh@sanger.ac.uk
COMMENT Constructed at the Institute for Genomic Research (TIGR),
Rockville, MD. Genomic DNA isolated from a cloned population of
Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared
to give a tight size distribution (
4 kb). The v + i method used for the library construction is
described in detail in Smith, H. and Venter, J.C. (Making small
insert libraries for whole genome shotgun sequencing projects. In
Genome Sequencing: A Practical Approach, eds. M. Vaubin and B.
Barrell, Oxford University Press, 1999).
Email: nelsayed@tigr.org
Details of T. brucei sequencing at the Sanger Centre are available
at http://www.sanger.ac.uk/projects/T_brucei/.
Location/Qualifiers
1..24
/organism="Trypanosoma brucei"
/strain="TREU927"
/db_xref="taxon:5691"
/clone="206g04"
BASE COUNT 8 a 8 c 5 g 3 t
ORIGIN
Query Match 100.0%; Score 7; DB 13; Length 24;
Best Local Similarity 100.0%; Pred. No. 3.7e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 tgttga 7
Db 11 TGTGGA 5

RESULT 15
LOCUS AI074857 25 bp mRNA EST 24-SEP-1998
DEFINITION oy27b11.s1 Soares-senescent_fibroblasts_NbHSP Homo sapiens cDNA
clone IMAGE:1667037 3' similar to TR:064507 064507 SERINE 1 ULTRA
HIGH SULFUR PROTEIN. ; contains PTR5 repetitive element ;, mRNA
sequence.
ACCESSION AI074857
VERSION AI074857.1 GI:3401501
KEYWORDS EST.
SOURCE EST.
ORGANISM Homo sapiens
human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

REFERENCE 1 (bases 1 to 25)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps@mail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Trace considered overall poor quality
Insert Length: 2304 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1..25
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1667037"
/tissue="Soares-senescent_fibroblasts_NbHSP"
/issue_type="senescent_fibroblast"
/lab_host="PH109 (ampicillin resistant)"
/notes="Vector: pT73D (Pharmacia) with a modified
polylinker V-type: phagemid; Site_1: Not I; Site_2: Eco RI
; 1st strand cDNA was primed with a Not I - oligo(dT)
primer [5'
TCCTACCAATCGAAGTGGAGCGCCGATTTTTTTTTTTTTTTT 3']
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M. Fatima Bonaldo."
BASE COUNT 1 a 4 c 9 g 11 t
ORIGIN
Query Match 100.0%; Score 7; DB 10; Length 25;
Best Local Similarity 100.0%; Pred. No. 3.7e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 tgttga 7
Db 2 TGTGGA 8

```

Search completed: December 15, 2001, 02:33:41
 Job time: 4981 sec

Mon Dec 17 07:48:29 2001

us-09-380-826a-4.rst

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 03:19:53 : Search time 2725.73 Seconds
(without alignments)
42.367 Million cell updates/sec

Title: US-09-380-826a-5
Perfect score: 7
Sequence: 1 ttgata 7

Scoring table: IDENTITY-NUC
Gapop 10.0, Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues
Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl: *
1: gb_da: *
2: gb_htg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vi: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_om: *
20: em_or: *
21: em_ov: *
22: em_pat: *
23: em_ph: *
24: em_pl: *
25: em_ro: *
26: em_sts: *
27: em_sy: *
28: em_un: *
29: em_vi: *
30: em_htgo_hum: *
31: em_htgo_inv: *
32: em_htgo_rod: *
33: em_htg_hum: *
34: em_htg_inv: *
35: em_htg_rod: *
36: em_htg_other: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	7	100.0	11	6 120741	120741 Sequence 6
C 2	7	100.0	11	6 120742	120742 Sequence 7
C 3	7	100.0	13	6 AR082373	AR082373 Sequence
C 4	7	100.0	13	6 AR120915	AR120915 Sequence
C 5	7	100.0	13	6 152063	152063 Sequence 5
C 6	7	100.0	13	6 178419	178419 Sequence 21
C 7	7	100.0	14	6 A88596	A88596 Sequence 74
C 8	7	100.0	14	6 A90563	A90563 Sequence 74
C 9	7	100.0	14	6 AR082368	AR082368 Sequence
C 10	7	100.0	14	6 AR120910	AR120910 Sequence
C 11	7	100.0	14	6 AX007639	AX007639 Sequence
C 12	7	100.0	14	6 178414	178414 Sequence 21
C 13	7	100.0	15	6 A35660	A35660 Synthetic h
C 14	7	100.0	15	6 AR045297	AR045297 Sequence
C 15	7	100.0	15	6 AR071525	AR071525 Sequence
C 16	7	100.0	15	6 AR131688	AR131688 Sequence
C 17	7	100.0	15	6 AR131689	AR131689 Sequence
C 18	7	100.0	15	6 AR131690	AR131690 Sequence
C 19	7	100.0	15	6 AR131691	AR131691 Sequence
C 20	7	100.0	15	6 AR133701	AR133701 Sequence
C 21	7	100.0	15	6 AR133702	AR133702 Sequence
C 22	7	100.0	15	6 AR133703	AR133703 Sequence
C 23	7	100.0	15	6 AR133704	AR133704 Sequence
C 24	7	100.0	15	6 AR133894	AR133894 Sequence
C 25	7	100.0	15	6 AR133895	AR133895 Sequence
C 26	7	100.0	15	6 AR133896	AR133896 Sequence
C 27	7	100.0	15	6 AX007638	AX007638 Sequence
C 28	7	100.0	15	6 AX007639	AX007639 Sequence
C 29	7	100.0	15	6 152349	152349 Sequence 90
C 30	7	100.0	15	6 161575	161575 Sequence 12
C 31	7	100.0	15	6 161576	161576 Sequence 13
C 32	7	100.0	15	6 161577	161577 Sequence 13
C 33	7	100.0	15	6 177301	177301 Sequence 8
C 34	7	100.0	15	6 177302	177302 Sequence 9
C 35	7	100.0	15	6 177303	177303 Sequence 10
C 36	7	100.0	15	6 177948	177948 Sequence 65
C 37	7	100.0	15	6 177949	177949 Sequence 65
C 38	7	100.0	16	6 A64456	A64456 Sequence 7
C 39	7	100.0	16	6 A64457	A64457 Sequence 8
C 40	7	100.0	16	6 A88876	A88876 Sequence 10
C 41	7	100.0	16	6 AR141309	AR141309 Sequence
C 42	7	100.0	16	6 AR141310	AR141310 Sequence
C 43	7	100.0	16	6 AX007634	AX007634 Sequence
C 44	7	100.0	16	6 AX007635	AX007635 Sequence
C 45	7	100.0	16	6 AX007637	AX007637 Sequence

ALIGNMENTS

RESULT	1	PAT	07-OCT-1996
120741/c	120741	11 bp	DNA
LOCUS	Sequence 6 from patent US 5516634.		
DEFINITION	120741		
ACCESSION	120741		
VERSION	120741.1		GT:1601096
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 11)		
AUTHORS	Newman, P.J. and Santoso, S.S.		
TITLE	Molecular basis of the human platelet bra/drba alloantigen system and applications thereof		
JOURNAL	Patent: US 5516634-A 6 14-MAY-1996;		
FEATURES	Location/Qualifiers		
SOURCE	1..11		
BASE COUNT	5 a 1 c 3 g 2 t		
ORIGIN	/organism="unknown"		

Query Match 100.0%; Score 7; DB 6; Length 11;
Best Local Similarity 100.0%; Pred. No. 5.8e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttgtgata 7
|||||||
DB 7 TTTGATA 1

RESULT 2
LOCUS I20742 11 bp DNA PAT 07-OCT-1996
DEFINITION Sequence 7 from patent US 5516634.
ACCESSION I20742
VERSION I20742.1 GI:1601097
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 11)
AUTHORS Newman,P.J. and Santoso,S.S.
TITLE Molecular basis of the human platelet bra/brb alloantigen system
and applications thereof
JOURNAL Patent: US 5516634-A 7 14-MAY-1996;
FEATURES Location/Qualifiers
source 1..11
/organism="unknown"
BASE COUNT 6 a 1 c 2 g 2 t
ORIGIN

Query Match 100.0%; Score 7; DB 6; Length 11;
Best Local Similarity 100.0%; Pred. No. 5.8e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttgtgata 7
|||||||
DB 7 TTTGATA 1

RESULT 3
LOCUS AR082373 13 bp DNA PAT 31-AUG-2000
DEFINITION Sequence 217 from patent US 5972704.
ACCESSION AR082373
VERSION AR082373.1 GI:10009099
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 13)
AUTHORS Draper,K.G., Chowrira,B., McSwiggen,J., Stinchcomb,D.T. and
Thompson,J.D.
TITLE HIV nef targeted ribozymes
JOURNAL Patent: US 5972704-A 217 26-OCT-1999;
FEATURES Location/Qualifiers
source 1..13
/organism="unknown"
BASE COUNT 5 a 4 c 1 g 3 t
ORIGIN

Query Match 100.0%; Score 7; DB 6; Length 13;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttgtgata 7
|||||||
DB 10 TTTGATA 4

RESULT 4
LOCUS AR120915 13 bp DNA PAT 16-MAY-2001
DEFINITION Sequence 217 from patent US 6159692.
ACCESSION AR120915
VERSION AR120915.1 GI:14104491
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 13)
AUTHORS Draper,K.G., Chowrira,B., McSwiggen,J., Stinchcomb,D.T. and
Thompson,J.D.
TITLE Method and reagent for inhibiting human immunodeficiency virus
replication
JOURNAL Patent: US 6159692-A 217 12-DEC-2000;
FEATURES Location/Qualifiers
source 1..13
/organism="unknown"
BASE COUNT 5 a 4 c 1 g 3 t
ORIGIN

Query Match 100.0%; Score 7; DB 6; Length 13;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttgtgata 7
|||||||
DB 10 TTTGATA 4

RESULT 5
LOCUS I52063 13 bp DNA PAT 07-OCT-1997
DEFINITION Sequence 5 from patent US 5646020.
ACCESSION I52063
VERSION I52063.1 GI:2473264
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 13)
AUTHORS Swiggen,J.A. and Mamone,J.Anthony.
TITLE Hammerhead ribozymes for preferred targets
JOURNAL Patent: US 5646020-A 5 08-JUL-1997;
FEATURES Location/Qualifiers
source 1..13
/organism="unknown"
BASE COUNT 5 a 4 c 1 g 3 t
ORIGIN

Query Match 100.0%; Score 7; DB 6; Length 13;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttgtgata 7
|||||||
DB 10 TTTGATA 4

RESULT 6
LOCUS I78419 13 bp DNA PAT 03-APR-1998
DEFINITION Sequence 217 from patent US 5693535.
ACCESSION I78419
VERSION I78419.1 GI:3014573
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 13)

AUTHORS Draper,K.G., Chowrira,B., McSwiggen,J., Stinchcomb,D.T. and

Thompson,J.D.

TITLE HIV targeted ribozymes

JOURNAL Patent: US 5693535-A 217 02-DEC-1997;

FEATURES Location/Qualifiers

source 1..13

BASE COUNT 5 a 4 c 1 g 3 t

ORIGIN

Query Match

Best Local Similarity 100.0%; Score 7; DB 6; Length 13;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttgata 7

|||||

Db 10 TTTGATA 4

RESULT 7

LOCUS A88596 14 bp DNA

DEFINITION Sequence 744 from Patent WO9833904.

ACCESSION A88596

VERSION A88596.1 GI:6737166

KEYWORDS

SOURCE

ORGANISM

REFERENCE 1 (bases 1 to 14)

AUTHORS Brysch,W. and Schlingensiepen,K.

TITLE AN ANTISENSE OLIGONUCLEOTIDE PREPARATION METHOD

JOURNAL Patent: WO 9833904-A 744 06-AUG-1998;

FEATURES BIOGOSTIK GES (DE); BRYSCH WOLFGANG (DE)

source 1..14

Location/Qualifiers

/db_xref="taxon:32644"

BASE COUNT 5 a 1 c 1 g 7 t

ORIGIN

Query Match

Best Local Similarity 100.0%; Score 7; DB 6; Length 14;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttgata 7

|||||

Db 1 TTTGATA 7

RESULT 8

LOCUS A90563 14 bp DNA

DEFINITION Sequence 744 from Patent EP0856579.

ACCESSION A90563

VERSION A90563.1 GI:6739079

KEYWORDS

SOURCE

ORGANISM

REFERENCE 1 (bases 1 to 14)

AUTHORS Brysch,W.D. and Schlingensiepen,K.D.

TITLE An antisense oligonucleotide preparation method

JOURNAL Patent: EP 0856579-A 744 05-AUG-1998;

FEATURES BIOGOSTIK GES (DE)

source 1..14

Location/Qualifiers

/db_xref="taxon:32644"

BASE COUNT 5 a 1 c 1 g 7 t

ORIGIN

Query Match

Best Local Similarity 100.0%; Score 7; DB 6; Length 14;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttgata 7

|||||

Db 1 TTTGATA 7

RESULT 9

LOCUS AR082368 14 bp DNA

DEFINITION Sequence 212 from patent US 5972704.

ACCESSION AR082368

VERSION AR082368.1 GI:10009094

KEYWORDS

SOURCE

ORGANISM

REFERENCE 1 (bases 1 to 14)

AUTHORS Draper,K.G., Chowrira,B., McSwiggen,J., Stinchcomb,D.T. and

TITLE HIV nef targeted ribozymes

JOURNAL Patent: US 5972704-A 212 26-OCT-1999;

FEATURES Location/Qualifiers

source 1..14

/db_xref="taxon:32644"

BASE COUNT 5 a 1 c 4 g 4 t

ORIGIN

Query Match

Best Local Similarity 100.0%; Score 7; DB 6; Length 14;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttgata 7

|||||

Db 3 TTTGATA 9

RESULT 10

LOCUS ARI20910 14 bp DNA

DEFINITION Sequence 212 from patent US 6159692.

ACCESSION ARI20910

VERSION ARI20910.1 GI:14104486

KEYWORDS

SOURCE

ORGANISM

REFERENCE 1 (bases 1 to 14)

AUTHORS Draper,K.G., Chowrira,B., McSwiggen,J., Stinchcomb,D.T. and

TITLE Method and reagent for inhibiting human immunodeficiency virus

JOURNAL Patent: US 6159692-A 212 12-DEC-2000;

FEATURES Location/Qualifiers

source 1..14

/db_xref="taxon:32644"

BASE COUNT 5 a 1 c 4 g 4 t

ORIGIN

Query Match

Best Local Similarity 100.0%; Score 7; DB 6; Length 14;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttgata 7

|||||

Db 3 TTTGATA 9

RESULT	11					
AX007639/c						
LOCUS	AX007639	14 bp	DNA	PAT		06-SEP-2000
DEFINITION	Sequence 181 from Patent WO967428.					
ACCESSION	AX007639					
VERSION	AX007639.1	GI:9995336				
KEYWORDS						
SOURCE	Aids-associated retrovirus.					
ORGANISM	Aids-associated retrovirus viruses; Retroid viruses; Retroviriidae. 1 (bases 1 to 14) Stuyver,L.					
REFERENCE	Method for detection of drug-selected mutations in the hiv protease gene					
AUTHORS	Patent: WO 9967428-A 181 29-DCC-1999:					
TITLE	INOGENETICS NV (BE); STUYVER LIEVEN (BE)					
JOURNAL	Location/Qualifiers					
FEATURES	source					
	1..14	/organism="Aids-associated retrovirus"				
		/db_xref="taxon:11966"				
BASE COUNT	5 a	1 c	2 g		6 t	
ORIGIN						

Query Match	100.0%;	Score 7;	DB 6;	Length 14;
Best Local Similarity	100.0%;	Pred. No. 5,7+05;		
Matches	7;	Conservative	0;	Mismatches 0;
				Indels 0;
				Gaps 0;
QY	1	tttgata	7	
Db	11	TTTGATA	5	

RESULT		12				
Locus	178A14				PAT	03-APR-1998
DEFINITION	Sequence	14 bp	DNA			
ACCESSION	178A14	212 from patent US 5693535.				
VERSION	178A14.1	G1:3014568				
KEYWORDS	.					
SOURCE	Unknown.					
ORGANISM	Unclassified.					
REFERENCE	1 (bases 1 to 14)					
AUTHORS	Draper,K.G., Chowlife,B., McSwiggen,J., Stinchcomb,D.T. and Thompson,J.D. Patent: US 5693535-A A 212 02 -DEC-1997; Location/Qualifiers					
TITLE	HIV targeted ribozymes					
JOURNAL						
FEATURES	. . .114					
Source	/organism="unknown"					
BASE COUNT	5 a	I c	4 g		4 t	
ORIGIN						

	Query Match	Score	DB	Length	Best Local Similarity	Pred.	No.	Mismatches	Indels	Gaps
Matches	7;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;	
QY	1 ttgtgata	7								
Db	3 ttttgata	9								

RESULT	13						
A35660/c	A35660	15 bp	DNA		PAT	02-DEC-1996	
LOCUS	A35660						
DEFINITION	Synthetic human IFN-alpha 2 gene oligo.						
ACCESSION	A35660						
VERSION	A35660.1	GI:1927042					
KEYWORDS							
SOURCE	synthetic construct.						

ORGANISM	synthetic construct artificial sequence.			
REFERENCE	1 (bases 1 to 15)			
AUTHORS	Camble, R. and Edge, M. D.			
TITLE	Analogous interferon polypeptides, process for their preparation and pharmaceutical compositions containing them			
JOURNAL	Patent: EP 0194006-A 105 10-SEP-1986; IMPERIAL CHEMICAL INDUSTRIES PLC			
FEATURES	Location/Qualifiers			
source	1..15			
	/organism="synthetic construct"			
	/db_xref="taxon:32630"			
BASE COUNT	5 a	4 c	2 g	4 t
ORIGIN				

Query Match	100.0%;	Score 7;	DB 6;	Length 15;
Best Local Similarity	100.0%;	Pred. No. 5.6e+05;		
Matches	7;	Conservative	0;	Mismatches 0;
				Indels 0;
				Gaps 0;
QY	1 ttgata	7		
Db	12 ttgata	6		

RESULT	14
AR045297/c	
LOCUS	AR045297 15 bp DNA PAT 29-SEP-1999
DEFINITION	Sequence from patent US 5817796.
ACCESSION	AR045297
VERSION	AR045297.1 GI:5966762
KEYWORDS	.
SOURCE	. Unknown.
ORGANISM	Unknown.
REFERENCE	Unclassified.
AUTHORS	1 (bases 1 to 15)
TITLE	Stluchcomb,D.T., Draper,K., McSwiggen,J. and Jarvis,T. C-mwb ribozymes having 2'-5'-linked adenylyate residues Patent: Us 5817796-A 90 06-Oct-1998;
JOURNAL	Location/Qualifiers
FEATURES	1..15
source	/organism="unknown"
BASE COUNT	6 a 3 c 3 g 3 t
ORIGIN	

Query Match	100.0%	Score 7;	DB 6;	length 15;
Best local Similarity	100.0%	Pred. No. 5.6e+05;		
Matches	7;	Conservative	0;	Mismatches 0;
				Indels 0;
				Gaps 0;
QY	1 ttgata 7			
Db	11 ttgata 5			

RESULT	15			
LOCUS	AR071525/c			
DEFINITION	Sequence	25	bp	DNA
ACCESSION	AR071525			
VERSION	AR071525.1			
KEYWORDS	GI:7222413			
SOURCE	Unknown.			
ORGANISM	Unknown.			
REFERENCE	Unclassified.			
AUTHORS	1 (bases 1 to 15)			
TITLE	Chao, Y.			
JOURNAL	H2-1 virus persistence-associated-gene 1 (PAG1) promoter uses			
FEATURES	therefor, and compositions containing same or products therefrom			
	Patent: US 5911987-A 25 15-JUN-1999;			
	Location/Qualifiers			
	1..15			
	/organism="unknown"			

BASE COUNT 9 a 2 c 2 g 2 t
ORIGIN

Query Match 100.0%; Score 7; DB 6; Length 15;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ttgata 7
|||
Db 12 TTGATA 6

Search completed: December 15, 2001, 03:19:53
Job time: 7683 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 03:29:44 : Search time 401.91 Seconds
(Without alignments)
14.932 Million cell updates/sec

Title: US-09-380-826a-5

Perfect score: 7

Sequence: 1 ttgata 7

Scoring table: IDENTITY_NUC
Gapop 10.0 , Capext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: N_Geneseq_1101:*
2: /SIDS2/gcgdata/geneseq/geneseq/NA1980.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseq/NA1981.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseq/NA1982.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseq/NA1983.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseq/NA1984.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseq/NA1985.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseq/NA1986.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseq/NA1987.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseq/NA1988.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseq/NA1989.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseq/NA1990.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseq/NA1991.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseq/NA1992.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseq/NA1993.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseq/NA1996.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	7	100.0	19	AAV58899
2	7	100.0	10	AAZ78577
3	7	100.0	10	AAZ60864
4	7	100.0	10	AAZ82567
5	7	100.0	10	AAZ82768
6	7	100.0	10	AAZ83117
7	7	100.0	10	AAZ83548
8	7	100.0	10	AAZ83548
9	7	100.0	10	AAZ83092
10	7	100.0	10	AAZ83092
11	7	100.0	10	AAZ83092
12	7	100.0	10	AAZ83092
13	7	100.0	10	AAZ83092
14	7	100.0	10	AAZ83092
15	7	100.0	10	AAZ83092
16	7	100.0	10	AAZ83092
17	7	100.0	10	AAZ83092
18	7	100.0	10	AAZ83092
19	7	100.0	10	AAZ83092
20	7	100.0	10	AAZ83092
21	7	100.0	10	AAZ83092
22	7	100.0	10	AAZ83092

12	7	100.0	10	AAZ83092	Yeast NORF gene SA
13	7	100.0	10	AAZ83092	Yeast NORF gene SA
14	7	100.0	10	AAZ83092	Yeast NORF gene SA
15	7	100.0	10	AAZ83092	Yeast NORF gene SA
16	7	100.0	10	AAZ83092	Yeast NORF gene SA
17	7	100.0	10	AAZ83092	Yeast NORF gene SA
18	7	100.0	10	AAZ83092	Yeast NORF gene SA
19	7	100.0	10	AAZ83092	Yeast NORF gene SA
20	7	100.0	10	AAZ83092	Yeast NORF gene SA
21	7	100.0	10	AAZ83092	Yeast NORF gene SA
22	7	100.0	10	AAZ83092	Yeast NORF gene SA
23	7	100.0	10	AAZ83092	Yeast NORF gene SA
24	7	100.0	10	AAZ83092	Yeast NORF gene SA
25	7	100.0	10	AAZ83092	Yeast NORF gene SA
26	7	100.0	10	AAZ83092	Yeast NORF gene SA
27	7	100.0	10	AAZ83092	Yeast NORF gene SA
28	7	100.0	10	AAZ83092	Yeast NORF gene SA
29	7	100.0	10	AAZ83092	Yeast NORF gene SA
30	7	100.0	10	AAZ83092	Yeast NORF gene SA
31	7	100.0	10	AAZ83092	Yeast NORF gene SA
32	7	100.0	10	AAZ83092	Yeast NORF gene SA
33	7	100.0	10	AAZ83092	Yeast NORF gene SA
34	7	100.0	10	AAZ83092	Yeast NORF gene SA
35	7	100.0	10	AAZ83092	Yeast NORF gene SA
36	7	100.0	10	AAZ83092	Yeast NORF gene SA
37	7	100.0	10	AAZ83092	Yeast NORF gene SA
38	7	100.0	10	AAZ83092	Yeast NORF gene SA
39	7	100.0	10	AAZ83092	Yeast NORF gene SA
40	7	100.0	10	AAZ83092	Yeast NORF gene SA
41	7	100.0	10	AAZ83092	Yeast NORF gene SA
42	7	100.0	10	AAZ83092	Yeast NORF gene SA
43	7	100.0	10	AAZ83092	Yeast NORF gene SA
44	7	100.0	10	AAZ83092	Yeast NORF gene SA
45	7	100.0	10	AAZ83092	Yeast NORF gene SA

ALIGNMENTS

RESULT 1	
ID AAV58899 standard; DNA; 7 BP.	
AAV58899;	
20-JAN-1999 (first entry)	
Leptospira nucleotide sequence.	
Infection: pathogenic Leptospira; protective immunity; therapy;	
diagnosis; ss.	
Leptospira sp.	
WO9840099-A1.	
17-SEP-1998.	
06-MAR-1998; 98WO-AU00145.	
07-MAR-1997; 97AU-0005494.	
(AGRI-) AGRIC VICTORIA SERVICES PTY LTD.	
(PIGR-) PIG RES & DEV CORP.	
Chappel RJ;	
WPI; 1998-520791/44.	
New isolated pathogenic Leptospira bacterium - useful for, e.g	
developing products for conferring protective immunity, and for	
prophylactic or therapeutic treatment	

Query Match	Best Local Similarity	Score 7:	DB 19:	Length 7:
Matches 7:	Conservative 0:	Mismatches 0:	Indels 0:	Gaps 0:
QY 1 ttgata 7 				
Db 1 ttgata 7				
RESULT 2				
AAZ78577/C				
ID AAZ78577 standard; DNA: 10 BP.				
XX AAZ78577;				
XX AC				
XX DT 10-APR-2000 (first entry)				
XX DE Human dendritic cell SAGE tag, SEQ ID NO:1005.				
XX DE				
XX KW SAGE tag; serial analysis of gene expression; antigen-presenting cell; APC; monocyte-derived dendritic cell; differential gene expression; immunostimulatory cofactor; costimulatory factor; CTL; cytotoxic T-lymphocyte; tumour antigen; immunotherapy; anticancer; ss.				
KM				
XX				
OS Homo sapiens.				
XX				
PN W0965924-A2.				
XX				
PD 23-DEC-1999.				
PF 18-JUN-1999;	99WO-US13800.			
XX				
PR 19-JUN-1998;	98US-0089833.			
PR 19-JUN-1998;	98US-0089844.			
PR 19-JUN-1998;	98US-0089853.			
PR 19-JUN-1998;	98US-0089878.			
PR 19-JUN-1998;	98US-0089991.			
PR 19-JUN-1998;	98US-0089992.			
PR 19-JUN-1998;	98US-0089993.			
PR 19-JUN-1998;	98US-0089994.			
PR 19-JUN-1998;	98US-0089997.			
PR 19-JUN-1998;	98US-0089999.			
PR 19-JUN-1998;	98US-0090000.			
PR 19-JUN-1998;	98US-0090035.			
PR 19-JUN-1998;	98US-0090036.			
PR 19-JUN-1998;	98US-0090039.			
PR 19-JUN-1998;	98US-0090040.			
PR 19-JUN-1998;	98US-0090041.			
PR 19-JUN-1998;	98US-0090042.			
PR 19-JUN-1998;	98US-0090043.			
PR 19-JUN-1998;	98US-0090044.			
PR 19-JUN-1998;	98US-0090045.			
PR 19-JUN-1998;	98US-0090047.			
PR 19-JUN-1998;	98US-0090048.			
PR 19-JUN-1998;	98US-0090072.			
PR 19-JUN-1998;	98US-0090076.			
PR 19-JUN-1998;	98US-0090077.			
PR 19-JUN-1998;	98US-0090078.			
PR 19-JUN-1998;	98US-0090079.			

PR	19-JUN-1998;	9805-009080.
PR	08-DEC-1998;	9805-011715.
PA	(GENZ.) GENZYME CORP.	
PA	(ROBE/) ROBERTS B. L.	
PA	(SHAN/) SHANKARA S.	
PI	Roberts BL,	Shankara S;
DR	WPI: 2000-106077/09.	
PT	Isolated polynucleotides differentially expressed in antigen-presenting	
XX	cells, useful in gene vaccines against cancer -	
XX	Claim 1; Page 94; 130pp: English.	
PS	Sequences AA277573..279709 represent SAGE (serial analysis of gene	
XX	expression) tags used to identify mRNA transcripts encoding	
CC	immunostimulatory cofactor proteins which are preferentially or	
CC	differentially expressed in monocyte-derived dendritic cells compared	
CC	with monocytes. Some of the transcripts correspond to known genes or	
CC	ESTs (expressed sequence tags) which were previously unknown to be	
CC	preferentially or differentially expressed in dendritic cells, while	
CC	other transcripts correspond to novel genes. Antigen-presenting cell	
CC	(APC)-associated costimulatory factors play an important role in the	
CC	activation of the cytotoxic immune response, particularly against tumour	
CC	cells. Tumour antigen presentation via the MHC (major histocompatibility	
CC	complex) and subsequent recognition by T-cell receptors is alone	
CC	insufficient to activate a robust cytotoxic immune response that can	
CC	lyse the tumour cells, immunostimulatory cofactors also being required	
CC	for efficient activation of cytotoxic T-lymphocytes (CTLs). Nucleic acid	
CC	sequences identified using the SAGE tags have several potential uses.	
CC	They may be used in vaccines to induce an immune response, particularly	
CC	against a tumour antigen; to modulate the genotype of an APC; to screen	
CC	for agents that modulate expression of differentially expressed genes in	
CC	an APC; and as hybridisation probes/amplification primers for the	
CC	diagnosis, prognosis and monitoring of diseases related to abnormal	
CC	expression of these genes. Detection of the dendritic cell	
CC	differentially expressed genes, or of their encoded proteins, can be used	
CC	to identify cells as belonging to the monocyte lineage. Cells containing	
CC	these genes can be used in active immunotherapy (or to stimulate	
CC	production of a population of antigen-specific effector cells) and	
CC	vectors containing them are used in gene therapy. Co-administration of	
CC	tumour antigens and APC-associated costimulatory factors ensures adequate	
CC	antigen presentation to endogenous APCs and upregulates the APCs for the	
CC	presentation of co-stimulatory signals, migration to T cell-rich sites,	
CC	secretion of T cell growth factors and secretion of chemokines for	
CC	recruitment of immune effector cells.	
XX		
SQ	Sequence 10 BP; 6 A; 1 C; 1 G; 2 T; 0 other:	
XX		
XX	Query Match	100.0%; Score 7; DB 21; Length 10;
XX	Best Local Similarity	100.0%; Pred. No. 4.3e+04;
XX	Matches 7; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
OY	1 ttgata 7	
DB		
	9 TTGATA 3	
RESULT	3	
AAZ80864		
ID	AAZ80864 standard; DNA; 10 BP.	
XX		
AC	AAZ80864;	
DT	07-APR-2000 (first entry)	
DE	Metastatic breast tumour cell upregulated transcript tag #98.	
XX	Human; metastatic breast tumour tissue; breast cancer; tag; primer;	
KM	non-metastatic breast tumour tissue; gene therapy; anticancer;	

KM		antimetastatic; vaccine; diagnosis; ss.
XX		
OS	Homo sapiens.	
XX		
PN	M0965928-A2.	
XX		
PD	23-DEC-1999.	
XX		
PJ	18-JUN-1999;	99WO-US13647.
XX		
PR	19-JUN-1998;	98US-0089853.
PR	19-JUN-1998;	98US-0089997.
PR	19-JUN-1998;	98US-0090039.
PR	19-JUN-1998;	98US-0090040.
PR	19-JUN-1998;	98US-0090041.
PA	(GENZ) GENZYME CORP.	
PA	(ROBE/) ROBERTS B L.	
PA	(SHAN/) SHANKARA S.	
PI	Roberts BL,	Shankara S:
XX		
DR	WPI: 2000-106079/09.	
PT	Isolated polynucleotides differentially expressed between metastatic and non-metastatic breast cancer cells, useful for diagnosis, prevention and treatment of cancer -	
XX		
PS	Claim 1: Page 60; 21pp; English.	
XX		
CC	Az80767 to Az83941 represent tags corresponding to distinct transcripts that are preferentially transcribed in the metastatic breast tumour tissue (i.e. are upregulated in metastatic breast tumour cells).	
CC	Az83942 to Az86677 represent tags corresponding to distinct transcripts that are preferentially transcribed in the primary or non-metastatic breast tumour tissue (i.e. are downregulated in metastatic breast tumour cells). These transcripts can be used for diagnosis, prognosis, monitoring and treatment of breast cancer, particularly where metastatic. Diagnosis is by standard immunoassays or hybridisation/amplification reactions. Compounds that modulate expression of the transcripts are potentially useful for treatment of (metastatic) breast cancer, while promoters from the transcripts are used to direct expression, in selected cell types, of e.g. therapeutic genes (also ribozymes or antisense sequences), particularly an antigen encoding sequence for use in gene or cell-based vaccines. Polypeptides encoded by the transcripts are also useful in vaccines for diagnosing breast cancer and for raising specific antibodies (Ab). Ab are used to detect the polypeptides as therapeutic agents. Host cells that produce the polypeptides can be used to expand and isolate populations of educated, antigen-specific immune effector cells, e.g. cytotoxic T lymphocytes, and these used for adoptive immunotherapy.	
SQ	Sequence 10 BP; 3 A; 0 C; 1 G; 6 T; 0 other;	
OY	Query Match	100.0%; Score 7; DB 21; Length 10;
	Best Local Similarity	100.0%; Pred. No. 4.3e+04;
	Matches 7; Conservative	0; Mismatches 0; Indels 0; Gaps 0.
	1 ttgata 7	
Db	3 ttgtata 9	
RESULT 4		
ID	Aaz82567	
AAZ82567	standard; DNA; 10 BP.	
AC	AAZ82567;	
DT	07-APR-2000 (first entry)	
DE	Metastatic breast tumour cell upregulated transcript tag #1801.	

```

XX Human; metastatic breast tumour tissue; breast cancer; tag; primer;
KW non-metastatic breast tumour tissue; gene therapy; anticancer;
KW antimetastatic; vaccine; diagnosis; ss.
OS Homo sapiens.
XX
XX WO9965928-A2.
PM
XX
XX 23-DEC-1999.
PD
XX
XX 18-JUN-1999; 99WO-US13647.
PF
XX
PR 19-JUN-1998; 98US-0089853.
PR 19-JUN-1998; 98US-0089897.
PR 19-JUN-1998; 98US-0090039.
PR 19-JUN-1998; 98US-0090040.
PR 19-JUN-1998; 98US-0090041.
XX
XX (GENZ ) GENZYME CORP.
PA (ROBE) ROBERTS B L.
PA (SHAN/) SHANKARA S.
PI
XX
XX Roberts BL, Shankara S;
PI
XX
XX WPI: 2000-106079/09.
XX
XX
XX Isolated polynucleotides differentially expressed between metastatic
PT and non-metastatic breast cancer cells, useful for diagnosis,
PT prevention and treatment of cancer -
XX
XX Claim 1; Page 107; 219pp; English.
XX
XX
XX AA280767 to AA283941 represent tags corresponding to distinct
CC transcripts that are preferentially transcribed in the metastatic breast
CC tumour tissue (i.e. are upregulated in metastatic breast tumour cells).
CC AA283942 to AA286677 represent tags corresponding to distinct transcripts
CC that are preferentially transcribed in the primary or non-metastatic
CC breast tumour tissue (i.e. are downregulated in metastatic breast tumour
CC cells). These transcripts can be used for diagnosis, prognosis,
CC monitoring and treatment of breast cancer, particularly where metastatic.
CC diagnosis is by standard immunoassays or hybridisation/amplification
CC reactions. Compounds that modulate expression of the transcripts are
CC potentially useful for treatment of (metastatic) breast cancer, while
CC promoters from the transcripts are used to direct expression, in selected
CC cell types, of e.g. therapeutic genes (also ribozymes or antisense
CC sequences), particularly an antigen-encoding sequence for use in gene or
CC cell-based vaccines. Polypeptides encoded by the transcripts are also
CC useful in vaccines, for diagnosing breast cancer and for raising
CC specific antibodies (Ab). Ab are used to detect the polypeptides or as
CC therapeutic agents. Host cells that produce the polypeptides can be used
CC to expand and isolate populations of educated, antigen-specific immune
CC effector cells, e.g. cytotoxic T lymphocytes, and these used for
CC adoptive immunotherapy.
XX
XX
XX Sequence 10 BP; 3 A; 0 C; 2 G; 5 T; 0 other;
XX
XX
XX Query Match 100.0%; Score 7; DB 21; Length 10;
XX Best Local Similarity 100.0%; Pred. No. 4.3e+04;
XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0.0
XX
XX QY 1 ttgata 7
XX |||||
XX Db 1 ttgata 7
XX
XX
XX RESULT 5
XX ID AA282768 standard; DNA; 10 BP.
XX AC AA282768;
XX

```

DT 07-APR-2000 (first entry)
XX Metastatic breast tumour cell upregulated transcript tag #2002.
DE
XX
XX Human: metastatic breast tumour tissue; breast cancer; tag: primer;
KW non-metastatic breast tumour tissue; gene therapy; anticancer;
XX antimetastatic; vaccine; diagnosis; ss.
OS Homo sapiens.
XX
XX WO965928-A2.
PN
XX
XX 23-DEC-1999.
PD
XX
XX 18-JUN-1999; 99WO-US13647.
PF
XX
XX 19-JUN-1998; 98US-0089853.
PR 19-JUN-1998; 98US-0089997.
PR 19-JUN-1998; 98US-0090039.
PR 19-JUN-1998; 98US-0090040.
PR 19-JUN-1998; 98US-0090041.
XX
XX (GENZ) GENZYME CORP.
PA (ROBE/) ROBERTS B L.
PA (SHAN/) SHANKARA S.
XX
XX Roberts BL, Shankara S;
PI
XX
XX WPI: 2000-106079/09.
DR
XX
XX Isolated polynucleotides differentially expressed between metastatic
PT and non-metastatic breast cancer cells, useful for diagnosis,
PT prevention and treatment of cancer -
XX
XX
XX Claim 1; Page 113; 219pp; English.
PS
XX
XX AA280767 to AA283941 represent tags corresponding to distinct
CC transcripts that are preferentially transcribed in the metastatic breast
CC tumour tissue (i.e. are upregulated in metastatic breast tumour cells).
CC AA283942 to AA286677 represent tags corresponding to distinct transcripts
CC that are preferentially transcribed in the primary or non-metastatic
CC breast tumour tissue (i.e. are downregulated in metastatic breast tumour
CC cells). These transcripts can be used for diagnosis, prognosis,
CC monitoring and treatment of breast cancer, particularly where metastatic.
CC reactions. Compounds that modulate expression of the transcripts are
CC potentially useful for treatment of (metastatic) breast cancer, while
CC promoters from the transcripts are used to direct expression, in selected
CC cell types, of e.g. therapeutic genes (also ribozymes or antisense
CC sequences), particularly an antigen-encoding sequence for use in gene or
CC cell-based vaccines. Polypeptides encoded by the transcripts are also
CC useful in vaccines, for diagnosing breast cancer and for raising
CC specific antibodies (Ab). Ab are used to detect the polypeptides or as
CC therapeutic agents. Host cells that produce the polypeptides can be used
CC to expand and isolate populations of educated, antigen-specific immune
CC effector cells, e.g. cytotoxic T lymphocytes, and these used for
CC adoptive immunotherapy.
CC
XX
XX Sequence 10 BP; 4 A; 0 C; 2 G; 4 T; 0 other;
SQ

Query Match 100.0%; Score 7; DB 21; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.3e+04;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ttgtata 7
| | | | |
Db 2 ttgtata 8

RESULT 6
AA283117
ID AA283117 standard; DNA: 10 BP.

XX
XX AA283117;
AC
XX
XX 07-APR-2000 (first entry)
DT
XX
XX Metastatic breast tumour cell upregulated transcript tag #2351.
DE
XX
XX Human: metastatic breast tumour tissue; breast cancer; tag: primer;
KW non-metastatic breast tumour tissue; gene therapy; anticancer;
XX antimetastatic; vaccine; diagnosis; ss.
OS Homo sapiens.
XX
XX WO965928-A2.
PN
XX
XX 23-DEC-1999.
PD
XX
XX 18-JUN-1999; 99WO-US13647.
PF
XX
XX 19-JUN-1998; 98US-0089853.
PR 19-JUN-1998; 98US-0089997.
PR 19-JUN-1998; 98US-0090039.
PR 19-JUN-1998; 98US-0090040.
PR 19-JUN-1998; 98US-0090041.
XX
XX (GENZ) GENZYME CORP.
PA (ROBE/) ROBERTS B L.
PA (SHAN/) SHANKARA S.
XX
XX Roberts BL, Shankara S;
PI
XX
XX WPI: 2000-106079/09.
DR
XX
XX Isolated polynucleotides differentially expressed between metastatic
PT and non-metastatic breast cancer cells, useful for diagnosis,
PT prevention and treatment of cancer -
XX
XX
XX Claim 1; Page 122; 219pp; English.
PS
XX
XX AA280767 to AA283941 represent tags corresponding to distinct
CC transcripts that are preferentially transcribed in the metastatic breast
CC tumour tissue (i.e. are upregulated in metastatic breast tumour cells).
CC AA283942 to AA286677 represent tags corresponding to distinct transcripts
CC that are preferentially transcribed in the primary or non-metastatic
CC breast tumour tissue (i.e. are downregulated in metastatic breast tumour
CC cells). These transcripts can be used for diagnosis, prognosis,
CC monitoring and treatment of breast cancer, particularly where metastatic.
CC reactions. Compounds that modulate expression of the transcripts are
CC potentially useful for treatment of (metastatic) breast cancer, while
CC promoters from the transcripts are used to direct expression, in selected
CC cell types, of e.g. therapeutic genes (also ribozymes or antisense
CC sequences), particularly an antigen-encoding sequence for use in gene or
CC cell-based vaccines. Polypeptides encoded by the transcripts are also
CC useful in vaccines, for diagnosing breast cancer and for raising
CC specific antibodies (Ab). Ab are used to detect the polypeptides or as
CC therapeutic agents. Host cells that produce the polypeptides can be used
CC to expand and isolate populations of educated, antigen-specific immune
CC effector cells, e.g. cytotoxic T lymphocytes, and these used for
CC adoptive immunotherapy.
CC
XX
XX Sequence 10 BP; 4 A; 0 C; 1 G; 5 T; 0 other;
SQ

Query Match 100.0%; Score 7; DB 21; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.3e+04;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ttgtata 7
| | | | |
Db 1 ttgtata 7

RESULT 7
AAF34548/C
ID AAF34548 standard; DNA; 10 BP.
XX
AC AAF34548;
XX
DT 23-MAR-2001 (first entry)
XX
DE Yeast NORF gene SAGE tag oligonucleotide seq ID NO:1287.
XX
KW Yeast; Saccharomyces cerevisiae; characterisation; cell cycle; NORF;
KW not previously assigned open reading frame; nonannotated ORF; SAGE;
KW Serial analysis of gene expression; antifungal; tag; identification;
KW linker; PCR primer; ds.
XX
OS Saccharomyces cerevisiae.
XX
PN W0200077214-A2.
XX
PD 21-DEC-2000.
XX
PF 14-JUN-2000; 2000WO-US16223.
XX
PR 16-JUN-1999; 99US-0335032.
XX
PA (UYJO) UNIV JOHNS HOPKINS.
XX
PI Velculescu V, Vogelstein B, Kinzler K;
XX
DR WPI: 2001-061874/07.
XX
PT Yeast gene coding sequences comprising NORF genes with serial analysis
PT of gene expression (SAGE) tags, useful for studying, monitoring and
PT affecting phases of the cell cycle -
XX
XX Example; Page 46; 419pp; English.
XX
CC The present invention describes an isolated DNA molecule comprising a
CC coding sequence of a yeast gene selected from a group of 745 NORF (not
CC previously assigned open reading frame; or nonannotated ORF) genes
CC comprising a SAGE (serial analysis of gene expression) tag. Also
CC described are: (1) a method (M1) of using NORF genes to affect the cell
CC cycle comprising administering a NORF gene whose expression varies by at
CC least 10% between any two phases of the cell cycle selected from log
CC phase, S phase and G2/M; (2) a method (M2) for screening candidate
CC antifungal drugs comprising: (a) contacting a test substance with a
CC yeast cell; and (b) monitoring expression of a NORF gene whose
CC expression varies as in M1, where a test substance which modifies the
CC expression of the yeast gene is a candidate antifungal drug; (3) a method
CC (M3) for identifying human genes which are involved in cell cycle
CC progression comprising contacting human DNA with a probe which comprises
CC at least 10 contiguous nucleotides of a NORF gene whose expression varies
CC as in M1; and (4) a method (M4) for identifying a candidate drug as a
CC member of a class of drugs having a characteristic effect on gene
CC expression in a yeast cell comprising contacting a yeast cell with a
CC candidate drug and monitoring expression in the yeast cell of at least 1
CC NORF gene whose expression is affected by the class of drugs. The NORF
CC genes may be used to study, monitor and affect phases of the cell cycle,
CC the differentially expressed genes may be used as markers of phases of
CC the cell cycle. The methods may be used to identify candidate drugs which
CC affect the cell cycle and for identification of antifungal drugs.
CC AAF33268 to AAF44064 represent SAGE tags used in the exemplification of
CC the present invention. AAF33262 to AAF33267 represent linkers and PCR
CC primers used in the SAGE method, in the exemplification of the present
CC invention.
XX
SQ Sequence 10 BP; 7 A; 1 C; 0 G; 2 T; 0 other;

QY 1 ttctgata 7
Db 8 TTTGATA 2
RESULT 8
AAF35054/C
ID AAF35054 standard; DNA; 10 BP.
XX
AC AAF35054;
XX
DT 23-MAR-2001 (first entry)
XX
DE Yeast NORF gene SAGE tag oligonucleotide seq ID NO:1793.
XX
KW Yeast; Saccharomyces cerevisiae; characterisation; cell cycle; NORF;
KW not previously assigned open reading frame; nonannotated ORF; SAGE;
KW Serial analysis of gene expression; antifungal; tag; identification;
KW linker; PCR primer; ds.
XX
OS Saccharomyces cerevisiae.
XX
PN W0200077214-A2.
XX
PD 21-DEC-2000.
XX
PF 14-JUN-2000; 2000WO-US16223.
XX
PR 16-JUN-1999; 99US-0335032.
XX
PA (UYJO) UNIV JOHNS HOPKINS.
XX
PI Velculescu V, Vogelstein B, Kinzler K;
XX
DR WPI: 2001-061874/07.
XX
PT Yeast gene coding sequences comprising NORF genes with serial analysis
PT of gene expression (SAGE) tags, useful for studying, monitoring and
PT affecting phases of the cell cycle -
XX
XX Example; Page 64; 419pp; English.
XX
CC The present invention describes an isolated DNA molecule comprising a
CC coding sequence of a yeast gene selected from a group of 745 NORF (not
CC previously assigned open reading frame; or nonannotated ORF) genes
CC comprising a SAGE (serial analysis of gene expression) tag. Also
CC described are: (1) a method (M1) of using NORF genes to affect the cell
CC cycle comprising administering a NORF gene whose expression varies by at
CC least 10% between any two phases of the cell cycle selected from log
CC phase, S phase and G2/M; (2) a method (M2) for screening candidate
CC antifungal drugs comprising: (a) contacting a test substance with a
CC yeast cell; and (b) monitoring expression of a NORF gene whose
CC expression varies as in M1, where a test substance which modifies the
CC expression of the yeast gene is a candidate antifungal drug; (3) a method
CC (M3) for identifying human genes which are involved in cell cycle
CC progression comprising contacting human DNA with a probe which comprises
CC at least 10 contiguous nucleotides of a NORF gene whose expression varies
CC as in M1; and (4) a method (M4) for identifying a candidate drug as a
CC member of a class of drugs having a characteristic effect on gene
CC expression in a yeast cell comprising contacting a yeast cell with a
CC candidate drug and monitoring expression in the yeast cell of at least 1
CC NORF gene whose expression is affected by the class of drugs. The NORF
CC genes may be used to study, monitor and affect phases of the cell cycle,
CC the differentially expressed genes may be used as markers of phases of
CC the cell cycle. The methods may be used to identify candidate drugs which
CC affect the cell cycle and for identification of antifungal drugs.
CC AAF33268 to AAF44064 represent SAGE tags used in the exemplification of
CC the present invention. AAF33262 to AAF33267 represent linkers and PCR
CC primers used in the SAGE method, in the exemplification of the present
CC invention.
XX
SQ Sequence 10 BP; 6 A; 2 C; 0 G; 2 T; 0 other;

Query Match 100.0%; Score 7; DB 22; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.3e+04;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match 100.0%; Score 7; DB 22; Length 10;
 Best Local Similarity 100.0%; Pred. No. 4.3e+04;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttgata 7
 |||||
 Db 8 TTTGATA 2

RESULT 9
 AAF38092/c
 ID AAF38092 standard; DNA; 10 BP.

XX AAF38092;

DT 23-MAR-2001 (first entry)

XX Yeast NORF gene SAGE tag oligonucleotide SEQ ID NO:4831.

XX Yeast: Saccharomyces cerevisiae; characterisation: cell cycle: NORF;
 KW nor previously assigned open reading frame; nonannotated ORF; SAGE;
 KW serial analysis of gene expression; antifungal; tag; identification;
 KW linker; PCR primer; ds.

XX Saccharomyces cerevisiae.

XX WO20007214-A2.

XX 21-DEC-2000.

XX 14-JUN-2000; 2000WO-US16223.

XX 16-JUN-1999; 99US-0335032.

XX (UYJO) UNIV JOHNS HOPKINS.

XX Velculescu V, Vogelstein B, Kinzler K;

XX WPI: 2001-061874/07.

PT Yeast gene coding sequences comprising NORF genes with serial analysis
 of gene expression (SAGE) tags, useful for studying, monitoring and
 affecting phases of the cell cycle -

XX Example; Page 172; 419pp; English.

XX The present invention describes an isolated DNA molecule comprising a
 coding sequence of a yeast gene selected from a group of 745 NORF (not
 previously assigned open reading frame; or nonannotated ORF) genes
 comprising a SAGE (serial analysis of gene expression) tag. Also
 described are: (1) a method (M1) of using NORF genes to affect the cell
 cycle comprising administering a NORF gene whose expression varies by at
 least 10% between any two phases of the cell cycle selected from log
 phase, S phase and G2/M; (2) a method (M2) for screening candidate
 antifungal drugs comprising: (a) contacting a test substance with a
 yeast cell; and (b) monitoring expression of a NORF gene whose
 expression varies as in M1, where a test substance which modifies the
 expression of the yeast gene is a candidate antifungal drug; (3) a method
 (M3) for identifying human genes which are involved in cell cycle
 progression comprising contacting human DNA with a probe which comprises
 at least 10 contiguous nucleotides of a NORF gene whose expression varies
 as in M1; and (4) a method (M4) for identifying a candidate drug as a
 member of a class of drugs having a characteristic effect on gene
 expression in a yeast cell comprising contacting a yeast cell with a
 candidate drug and monitoring expression in the yeast cell of at least 1
 NORF gene whose expression is affected by the class of drugs. The NORF
 genes may be used to study, monitor and affect phases of the cell cycle,
 the differentially expressed genes may be used as markers of phases of
 the cell cycle. The methods may be used to identify candidate drugs which
 affect the cell cycle and for identification of antifungal drugs.
 CC AAF33268 to AAF44064 represent SAGE tags used in the exemplification of
 the present invention. AAF33262 to AAF33267 represent linkers and PCR

CC primers used in the SAGE method, in the exemplification of the present
 CC invention.
 XX
 CC Sequence 10 BP; 4 A; 2 C; 0 G; 4 T; 0 other;

Query Match 100.0%; Score 7; DB 22; Length 10;
 Best Local Similarity 100.0%; Pred. No. 4.3e+04;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttgata 7
 |||||
 Db 7 TTTGATA 1

RESULT 10
 AAF39829/c
 ID AAF39829 standard; DNA; 10 BP.

XX AAF39829;

DT 23-MAR-2001 (first entry)

XX Yeast NORF gene SAGE tag oligonucleotide SEQ ID NO:6568.

XX Yeast: Saccharomyces cerevisiae; characterisation: cell cycle: NORF;
 KW nor previously assigned open reading frame; nonannotated ORF; SAGE;
 KW serial analysis of gene expression; antifungal; tag; identification;
 KW linker; PCR primer; ds.

XX Saccharomyces cerevisiae.

XX WO20007214-A2.

XX 21-DEC-2000.

XX 14-JUN-2000; 2000WO-US16223.

XX 16-JUN-1999; 99US-0335032.

XX (UYJO) UNIV JOHNS HOPKINS.

XX Velculescu V, Vogelstein B, Kinzler K;

XX WPI: 2001-061874/07.

PT Yeast gene coding sequences comprising NORF genes with serial analysis
 of gene expression (SAGE) tags, useful for studying, monitoring and
 affecting phases of the cell cycle -

XX Example; Page 234; 419pp; English.

XX The present invention describes an isolated DNA molecule comprising a
 coding sequence of a yeast gene selected from a group of 745 NORF (not
 previously assigned open reading frame; or nonannotated ORF) genes
 comprising a SAGE (serial analysis of gene expression) tag. Also
 described are: (1) a method (M1) of using NORF genes to affect the cell
 cycle comprising administering a NORF gene whose expression varies by at
 least 10% between any two phases of the cell cycle selected from log
 phase, S phase and G2/M; (2) a method (M2) for screening candidate
 antifungal drugs comprising: (a) contacting a test substance with a
 yeast cell; and (b) monitoring expression of a NORF gene whose
 expression varies as in M1, where a test substance which modifies the
 expression of the yeast gene is a candidate antifungal drug; (3) a method
 (M3) for identifying human genes which are involved in cell cycle
 progression comprising contacting human DNA with a probe which comprises
 at least 10 contiguous nucleotides of a NORF gene whose expression varies
 as in M1; and (4) a method (M4) for identifying a candidate drug as a
 member of a class of drugs having a characteristic effect on gene
 expression in a yeast cell comprising contacting a yeast cell with a
 candidate drug and monitoring expression in the yeast cell of at least 1
 NORF gene whose expression is affected by the class of drugs. The NORF
 genes may be used to study, monitor and affect phases of the cell cycle,

CC the differentially expressed genes may be used as markers of phases of
CC the cell cycle. The methods may be used to identify candidate drugs which
CC affect the cell cycle and for identification of antifungal drugs.
CC AAF33268 to AAF44064 represent SAGE tags used in the exemplification of
CC the present invention. AAF33262 to AAF33267 represent linkers and PCR
CC primers used in the SAGE method, in the exemplification of the present
CC invention.
XX
XX
SO Sequence 10 BP; 5 A; 1 C; 1 G; 3 T; 0 other;

Query Match 100.0%; Score 7; DB 22; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.3e+04;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ttgata 7
Db 10 TTGATA 4

RESULT 11
AAFA0369/c
ID AAF40369 standard; DNA: 10 BP.
XX
XX AAF40369;
AC
DT 23-MAR-2001 (first entry)
XX
DE Yeast NORF gene SAGE tag oligonucleotide SEQ ID NO:7108.
XX
XX Yeast; Saccharomyces cerevisiae; characterisation: cell cycle; NORF;
KM nor previously assigned open reading frame; nonannotated ORF; SAGE;
KW serial analysis of gene expression; antifungal; tag; identification;
KM linker; PCR primer; ds.
XX
XX OS Saccharomyces cerevisiae.
XX PN WO200077214-A2.
XX PD 21-DEC-2000.
XX PF 14-JUN-2000; 2000WO-US16223.
XX PR 16-JUN-1999; 99US-0335032.
XX PA (UYJO) UNIV JOHNS HOPKINS.
XX PI Velculescu V, Vogelstein B, Kinzler K;
XX DR WPI; 2001-061874/07.
XX
XX PT Yeast gene coding sequences comprising NORF genes with serial analysis
PT of gene expression (SAGE) tags, useful for studying, monitoring and
PT affecting phases of the cell cycle -
XX
XX Example; Page 253; 419pp; English.
XX
XX The present invention describes an isolated DNA molecule comprising a
CC coding sequence of a yeast gene selected from a group of 745 NORF (not
CC previously assigned open reading frame; or nonannotated ORF) genes
CC comprising a SAGE (serial analysis of gene expression) tag. Also
CC described are: (1) a method (M1) of using NORF genes to affect the cell
CC cycle comprising administering a NORF gene whose expression varies by at
CC least 10% between any two phases of the cell cycle selected from log
CC phase, S phase and G2/M; (2) a method (M2) for screening candidate
CC antifungal drugs comprising: (a) contacting a test substance with a
CC yeast cell; and (b) monitoring expression of a NORF gene whose
CC expression varies as in M1, where a test substance which modifies the
CC expression of the yeast gene is a candidate antifungal drug; (3) a method
CC (M3) for identifying human genes which are involved in cell cycle
CC progression comprising contacting human DNA with a probe which comprises
CC at least 10 contiguous nucleotides of a NORF gene whose expression varies
CC as in M1; and (4) a method (M4) for identifying a candidate drug as a

CC member of a class of drugs having a characteristic effect on gene
CC expression in a yeast cell comprising contacting a yeast cell with a
CC candidate drug and monitoring expression in the yeast cell of at least 1
CC NORF gene whose expression is affected by the class of drugs. The NORF
CC genes may be used to study, monitor and affect phases of the cell cycle,
CC the differentially expressed genes may be used as markers of phases of
CC the cell cycle. The methods may be used to identify candidate drugs which
CC affect the cell cycle and for identification of antifungal drugs.
CC AAF33268 to AAF44064 represent SAGE tags used in the exemplification of
CC the present invention. AAF33262 to AAF33267 represent linkers and PCR
CC primers used in the SAGE method, in the exemplification of the present
CC invention.
XX
XX
SO Sequence 10 BP; 5 A; 1 C; 0 G; 4 T; 0 other;

Query Match 100.0%; Score 7; DB 22; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.3e+04;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ttgata 7
Db 10 TTGATA 4

RESULT 12
AAFA0389
ID AAF40389 standard; DNA: 10 BP.
XX
XX AAF40389;
AC
DT 23-MAR-2001 (first entry)
XX
DE Yeast NORF gene SAGE tag oligonucleotide SEQ ID NO:7128.
XX
XX Yeast; Saccharomyces cerevisiae; characterisation: cell cycle; NORF;
KM nor previously assigned open reading frame; nonannotated ORF; SAGE;
KW serial analysis of gene expression; antifungal; tag; identification;
KM linker; PCR primer; ds.
XX
XX OS Saccharomyces cerevisiae.
XX PN WO200077214-A2.
XX PD 21-DEC-2000.
XX PF 14-JUN-2000; 2000WO-US16223.
XX PR 16-JUN-1999; 99US-0335032.
XX PA (UYJO) UNIV JOHNS HOPKINS.
XX PI Velculescu V, Vogelstein B, Kinzler K;
XX DR WPI; 2001-061874/07.
XX
XX PT Yeast gene coding sequences comprising NORF genes with serial analysis
PT of gene expression (SAGE) tags, useful for studying, monitoring and
PT affecting phases of the cell cycle -
XX
XX Example; Page 254; 419pp; English.
XX
XX The present invention describes an isolated DNA molecule comprising a
CC coding sequence of a yeast gene selected from a group of 745 NORF (not
CC previously assigned open reading frame; or nonannotated ORF) genes
CC comprising a SAGE (serial analysis of gene expression) tag. Also
CC described are: (1) a method (M1) of using NORF genes to affect the cell
CC cycle comprising administering a NORF gene whose expression varies by at
CC least 10% between any two phases of the cell cycle selected from log
CC phase, S phase and G2/M; (2) a method (M2) for screening candidate
CC antifungal drugs comprising: (a) contacting a test substance with a
CC yeast cell; and (b) monitoring expression of a NORF gene whose
CC expression varies as in M1, where a test substance which modifies the

expression of the yeast gene is a candidate antifungal drug; (3) a method (M3) for identifying human genes which are involved in cell cycle progression comprising contacting human DNA with a probe which comprises at least 10 contiguous nucleotides of a NORF gene whose expression varies as in M1; and (4) a method (M4) for identifying a candidate drug as a member of a class of drugs having a characteristic effect on gene expression in a yeast cell comprising contacting a yeast cell with a candidate drug and monitoring expression in the yeast cell of at least 1 NORF gene whose expression is affected by the class of drugs. The NORF genes may be used to study, monitor and affect phases of the cell cycle, the differentially expressed genes may be used as markers of phases of the cell cycle. The methods may be used to identify candidate drugs which affect the cell cycle and for identification of antifungal drugs.

AAF33268 to AAF44064 represent SAGE tags used in the exemplification of the present invention. AAF33262 to AAF33267 represent linkers and PCR primers used in the SAGE method, in the exemplification of the present invention.

Sequence 10 BP; 3 A; 1 C; 1 G; 5 T; 0 other;

Query Match
Best Local Similarity 100.0%; Score 7; DB 22; Length 10;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 tttgata 7
|||||||
Db 2 tttgata 8

RESULT 13
AAF40414
ID AAF40414 standard; DNA; 10 BP.
XX
AC AAF40414;
XX
DE 23-MAR-2001 (first entry)
XX
DE Yeast NORF gene SAGE tag oligonucleotide SEQ ID NO:7153.
XX
XX Yeast: Saccharomyces cerevisiae; characterisation; cell cycle; NORF;
KW nor previously assigned open reading frame; nonannotated ORF; SAGE;
KW serial analysis of gene expression; antifungal; tag; identification;
KW linker; PCR primer; ds.
XX
OS Saccharomyces cerevisiae.
XX
PN WO200077214-A2.
PD 21-DEC-2000.
XX
PF 14-JUN-2000; 2000WO-US16223.
XX
PR 16-JUN-1999; 99US-0335032.
XX
PA (UYJO) UNIV JOHNS HOPKINS.
XX
PI Velculescu V, Vogelstein B, Kinzler K;
XX
DR WPI: 2001-061874/07.
XX
XX Yeast gene coding sequences comprising NORF genes with serial analysis
PT of gene expression (SAGE) tags, useful for studying, monitoring and
PT affecting phases of the cell cycle -
XX
PS Example: Page 255; 419pp; English.
XX
XX The present invention describes an isolated DNA molecule comprising a
CC coding sequence of a yeast gene selected from a group of 745 NORF (not
CC previously assigned open reading frame; or nonannotated ORF) genes
CC comprising a SAGE (serial analysis of gene expression) tag. Also
CC described are: (1) a method (M1) of using NORF genes to affect the cell
cycle comprising administering a NORF gene whose expression varies by at

least 10% between any two phases of the cell cycle selected from log phase, S phase and G2/M; (2) a method (M2) for screening candidate antifungal drugs comprising: (a) contacting a test substance with a yeast cell; and (b) monitoring expression of a NORF gene whose expression varies as in M1, where a test substance which modifies the expression of the yeast gene is a candidate antifungal drug; (3) a method (M3) for identifying human genes which are involved in cell cycle progression comprising contacting human DNA with a probe which comprises at least 10 contiguous nucleotides of a NORF gene whose expression varies as in M1; and (4) a method (M4) for identifying a candidate drug as a member of a class of drugs having a characteristic effect on gene expression in a yeast cell comprising contacting a yeast cell with a candidate drug and monitoring expression in the yeast cell of at least 1 NORF gene whose expression is affected by the class of drugs. The NORF genes may be used to study, monitor and affect phases of the cell cycle, the differentially expressed genes may be used as markers of phases of the cell cycle. The methods may be used to identify candidate drugs which affect the cell cycle and for identification of antifungal drugs.

AAF33268 to AAF44064 represent SAGE tags used in the exemplification of the present invention. AAF33262 to AAF33267 represent linkers and PCR primers used in the SAGE method, in the exemplification of the present invention.

Sequence 10 BP; 3 A; 1 C; 1 G; 5 T; 0 other;

Query Match
Best Local Similarity 100.0%; Score 7; DB 22; Length 10;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 tttgata 7
|||||||
Db 3 tttgata 9

RESULT 14
AAF40778
ID AAF40778 standard; DNA; 10 BP.
XX
AC AAF40778;
XX
DE 23-MAR-2001 (first entry)
XX
DE Yeast NORF gene SAGE tag oligonucleotide SEQ ID NO:7517.
XX
XX Yeast: Saccharomyces cerevisiae; characterisation; cell cycle; NORF;
KW nor previously assigned open reading frame; nonannotated ORF; SAGE;
KW serial analysis of gene expression; antifungal; tag; identification;
KW linker; PCR primer; ds.
XX
OS Saccharomyces cerevisiae.
XX
PN WO200077214-A2.
PD 21-DEC-2000.
XX
PF 14-JUN-2000; 2000WO-US16223.
XX
PR 16-JUN-1999; 99US-0335032.
XX
PA (UYJO) UNIV JOHNS HOPKINS.
XX
PI Velculescu V, Vogelstein B, Kinzler K;
XX
DR WPI: 2001-061874/07.
XX
XX Yeast gene coding sequences comprising NORF genes with serial analysis
PT of gene expression (SAGE) tags, useful for studying, monitoring and
PT affecting phases of the cell cycle -
XX
PS Example: Page 268; 419pp; English.
XX
XX The present invention describes an isolated DNA molecule comprising a

CC coding sequence of a yeast gene selected from a group of 745 NORF (not
 CC previously assigned open reading frame; or nonannotated ORF) genes
 CC comprising a SAGE (serial analysis of gene expression) tag. Also
 CC described are: (1) a method (M1) of using NORF genes to affect the cell
 CC cycle comprising administering a NORF gene whose expression varies by at
 CC least 10% between any two phases of the cell cycle selected from 1og
 CC phase, S phase and G2/M; (2) a method (M2) for screening candidate
 CC antifungal drugs comprising: (a) contacting a test substance with a
 CC yeast cell; and (b) monitoring expression of a NORF gene whose
 CC expression varies as in M1, where a test substance which modifies the
 CC expression of the yeast gene is a candidate antifungal drug; (3) a method
 CC (M3) for identifying human genes which are involved in cell cycle
 CC progression comprising contacting human DNA with a probe which comprises
 CC at least 10 contiguous nucleotides of a NORF gene whose expression varies
 CC as in M1; and (4) a method (M4) for identifying a candidate drug as a
 CC member of a class of drugs having a characteristic effect on gene
 CC expression in a yeast cell comprising contacting a yeast cell with a
 CC candidate drug and monitoring expression in the yeast cell of at least 1
 CC NORF gene whose expression is affected by the class of drugs. The NORF
 CC genes may be used to study, monitor and affect phases of the cell cycle,
 CC the differentially expressed genes may be used as markers of phases of
 CC the cell cycle. The methods may be used to identify candidate drugs which
 CC affect the cell cycle and for identification of antifungal drugs.
 CC AAF33268 to AAF44064 represent SAGE tags used in the exemplification of
 CC the present invention. AAF33262 to AAF33267 represent linkers and PCR
 CC primers used in the SAGE method, in the exemplification of the present
 CC invention.
 CC XX
 S0 Sequence 10 BP; 3 A; 1 C; 2 G; 4 T; 0 other;

Query Match 100.0%; Score 7; DB 22; Length 10;
 Best Local Similarity 100.0%; Pred. No. 4.3e+04;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttgata 7
 |||||
 Db 3 ttgata 9

RESULT 15
 AAF41077
 ID AAF41077 standard; DNA; 10 BP.
 XX
 AC AAF41077;
 XX
 DT 23-MAR-2001 (first entry)
 XX
 DE Yeast NORF gene SAGE tag oligonucleotide SEQ ID NO:7816.
 XX
 KW Yeast: *Saccharomyces cerevisiae*; Characterisation: cell cycle; NORF;
 KW not previously assigned open reading frame; nonannotated ORF; SAGE;
 KW serial analysis of gene expression; antifungal; tag; Identification;
 KW linker; PCR primer; ds.
 XX
 OS *Saccharomyces cerevisiae*.
 XX
 PN WO200077214-A2.
 XX
 PD 21-DEC-2000.
 XX
 PF 14-JUN-2000; 2000WO-US16223.
 XX
 PR 16-JUN-1999; 99US-0335032.
 XX
 PA (UJJO) UNIV JOHNS HOPKINS.
 XX
 PI Velculescu V, Vogelstein B, Kinzler K;
 XX
 DR WPI: 2001-061874/07.
 XX
 PT Yeast gene coding sequences comprising NORF genes with serial analysis
 PT of gene expression (SAGE) tags, useful for studying, monitoring and

PT affecting phases of the cell cycle -
 XX
 PS Example; Page 279; 419pp; English.

XX The present invention describes an isolated DNA molecule comprising a
 CC coding sequence of a yeast gene selected from a group of 745 NORF (not
 CC previously assigned open reading frame; or nonannotated ORF) genes
 CC comprising a SAGE (serial analysis of gene expression) tag. Also
 CC described are: (1) a method (M1) of using NORF genes to affect the cell
 CC cycle comprising administering a NORF gene whose expression varies by at
 CC least 10% between any two phases of the cell cycle selected from 1og
 CC phase, S phase and G2/M; (2) a method (M2) for screening candidate
 CC antifungal drugs comprising: (a) contacting a test substance with a
 CC yeast cell; and (b) monitoring expression of a NORF gene whose
 CC expression varies as in M1, where a test substance which modifies the
 CC expression of the yeast gene is a candidate antifungal drug; (3) a method
 CC (M3) for identifying human genes which are involved in cell cycle
 CC progression comprising contacting human DNA with a probe which comprises
 CC at least 10 contiguous nucleotides of a NORF gene whose expression varies
 CC as in M1; and (4) a method (M4) for identifying a candidate drug as a
 CC member of a class of drugs having a characteristic effect on gene
 CC expression in a yeast cell comprising contacting a yeast cell with a
 CC candidate drug and monitoring expression in the yeast cell of at least 1
 CC NORF gene whose expression is affected by the class of drugs. The NORF
 CC genes may be used to study, monitor and affect phases of the cell cycle,
 CC the differentially expressed genes may be used as markers of phases of
 CC the cell cycle. The methods may be used to identify candidate drugs which
 CC affect the cell cycle and for identification of antifungal drugs.
 CC AAF33268 to AAF44064 represent SAGE tags used in the exemplification of
 CC the present invention. AAF33262 to AAF33267 represent linkers and PCR
 CC primers used in the SAGE method, in the exemplification of the present
 CC invention.
 CC XX
 S0 Sequence 10 BP; 2 A; 2 C; 1 G; 5 T; 0 other;

Query Match 100.0%; Score 7; DB 22; Length 10;
 Best Local Similarity 100.0%; Pred. No. 4.3e+04;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttgata 7
 |||||
 Db 3 ttgata 9

Search completed: December 15, 2001, 03:29:44
 Job time: 7884 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 03:22:47 ; Search time 172.39 Seconds
(Without alignments)
9.196 Million cell updates/sec

Title: US-09-380-826a-5
Perfect score: 7
Sequence: 1 ttgtgata 7

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 11323899 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/1na/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/1na/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/1na/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/1na/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/1na/PCPUS.COMB.seq:*
6: /cgn2_6/ptodata/2/1na/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	7	100.0	11	1	US-08-086-634-6 Sequence 6, Appli
C 2	7	100.0	11	1	US-08-086-634-7 Sequence 7, Appli
C 3	7	100.0	12	4	US-09-042-353-4 Sequence 4, Appli
C 4	7	100.0	12	4	US-08-758-617A-272 Sequence 5, Appli
C 5	7	100.0	13	1	US-08-166-664-5 Sequence 217, App
C 6	7	100.0	13	1	US-08-271-880A-217 Sequence 217, App
C 7	7	100.0	13	2	US-08-910-408-217 Sequence 217, App
C 8	7	100.0	13	3	US-09-249-215-217 Sequence 212, App
C 9	7	100.0	14	1	US-08-271-880A-212 Sequence 212, App
C 10	7	100.0	14	2	US-08-910-408-212 Sequence 212, App
C 11	7	100.0	14	3	US-09-249-215-212 Sequence 212, App
C 12	7	100.0	15	1	US-08-373-124A-90 Sequence 129, App
C 13	7	100.0	15	1	US-08-291-932A-129 Sequence 130, App
C 14	7	100.0	15	1	US-08-291-932A-130 Sequence 131, App
C 15	7	100.0	15	1	US-08-291-932A-131 Sequence 8, Appli
C 16	7	100.0	15	1	US-08-334-847-8 Sequence 9, Appli
C 17	7	100.0	15	1	US-08-334-847-9 Sequence 10, Appli
C 18	7	100.0	15	1	US-08-334-847-10 Sequence 655, App
C 19	7	100.0	15	1	US-08-334-847-655 Sequence 656, App
C 20	7	100.0	15	1	US-08-334-847-656 Sequence 90, Appli
C 21	7	100.0	15	2	US-08-585-684B-113 Sequence 113, App
C 22	7	100.0	15	2	US-08-585-684B-114 Sequence 114, App
C 23	7	100.0	15	2	US-08-585-684B-115 Sequence 115, App
C 24	7	100.0	15	2	US-08-585-684B-116 Sequence 116, App
C 25	7	100.0	15	2	US-08-585-684B-117 Sequence 2126, Ap
C 26	7	100.0	15	2	US-08-585-684B-2126 Sequence 2127, Ap
C 27	7	100.0	15	2	US-08-585-684B-2127

C 28	7	100.0	15	2	US-08-585-684B-2128 Sequence 2128, Ap
C 29	7	100.0	15	2	US-08-585-684B-2129 Sequence 2129, Ap
C 30	7	100.0	15	2	US-08-585-684B-2319 Sequence 2319, Ap
C 31	7	100.0	15	2	US-08-585-684B-2320 Sequence 2320, Ap
C 32	7	100.0	15	2	US-08-585-684B-2321 Sequence 2321, Ap
C 33	7	100.0	15	2	US-08-585-684B-2321 Sequence 25, Appli
C 34	7	100.0	15	2	US-08-634-350-25 Sequence 113, App
C 35	7	100.0	15	4	US-09-038-073-113 Sequence 114, App
C 36	7	100.0	15	4	US-09-038-073-114 Sequence 115, App
C 37	7	100.0	15	4	US-09-038-073-115 Sequence 116, App
C 38	7	100.0	15	4	US-09-038-073-116 Sequence 2126, Ap
C 39	7	100.0	15	4	US-09-038-073-2127 Sequence 2127, Ap
C 40	7	100.0	15	4	US-09-038-073-2128 Sequence 2128, Ap
C 41	7	100.0	15	4	US-09-038-073-2129 Sequence 2319, Ap
C 42	7	100.0	15	4	US-09-038-073-2319 Sequence 2320, Ap
C 43	7	100.0	15	4	US-09-038-073-2320 Sequence 2321, Ap
C 44	7	100.0	15	4	US-09-038-073-2321 Sequence 8, Appli
C 45	7	100.0	16	3	US-08-930-917A-8

ALIGNMENTS

RESULT 1
US-08-086-634-6/c
Sequence 6, Application US/08086634
Patent No. 5516634
GENERAL INFORMATION:
APPLICANT: NEWMAN, Peter J.
APPLICANT: SANTOSO, Sento Saleh
TITLE OF INVENTION: Molecular Basis of the
TITLE OF INVENTION: Human Platelet Brb/Brb Alloantigen System and
TITLE OF INVENTION: Applications Thereof
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Foley & Lardner
STREET: P. O. Box 1497
CITY: Madison
STATE: Wisconsin
COUNTRY: USA
ZIP: 53701-1497
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect, Version 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/086,634
FILING DATE: 30-June-1993
ATTORNEY/AGENT INFORMATION:
NAME: Scanlon, William J.
REGISTRATION NUMBER: 30,136
REFERENCE/DOCKET NUMBER: 30383/132
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 258-4284
TELEFAX: (608) 258-4258
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA to mRNA
US-08-086-634-6
Query Match 100.0%; Score 7; DB 1; Length 11;
Best Local Similarity 100.0%; Pred. No. 9.1e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 ttgtgata 7
Db 7 TTTGATA 1

```

RESULT 2
US-08-086-634-7/C
Sequence 7, Application US/08086634
Patent No. 5516634
GENERAL INFORMATION:
APPLICANT: NEWMAN, Peter J.
APPLICANT: SANTOSO, Senot Saleh
TITLE OF INVENTION: Molecular Basis of the
TITLE OF INVENTION: Human Platelet Bra/Brb Alloantigen System and
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: P. O. Box 1497
CITY: Madison
STATE: Wisconsin
COUNTRY: USA
ZIP: 53701-1497
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect, Version 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/086,634
FILING DATE: 30-June-1993
ATTORNEY/AGENT INFORMATION:
NAME: Scanlon, William J.
REGISTRATION NUMBER: 30,136
REFERENCE/DOCKET NUMBER: 30383/132
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 258-4284
TELEFAX: (608) 258-4258
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
US-08-086-634-7

Query Match 100.0%; Score 7; DB 1; Length 11;
Best Local Similarity 100.0%; Pred. No. 9.1e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ttgtata 7
Db 7 TTTGATA 1

RESULT 3
US-09-042-353-4
Sequence 4, Application US/09042353
Patent No. 6255458
GENERAL INFORMATION:
APPLICANT: Lonberg, Nils
APPLICANT: KAY, Robert M.
TITLE OF INVENTION: Transgenic No. 6255458-Human Animals for
TITLE OF INVENTION: Producing Heterologous Antibodies
NUMBER OF SEQUENCES: 421
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

```

```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/042,353
FILING DATE: 13-MAR-1998
CLASSIFICATION: 800
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/810,279
FILING DATE: 17-DEC-1991
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/853,408
FILING DATE: 18-MAR-1992
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/904,068
FILING DATE: 23-JUN-1992
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/990,860
FILING DATE: 16-DEC-1992
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/053,131
FILING DATE: 26-APR-1993
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/096,762
FILING DATE: 22-JUL-1993
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/155,301
FILING DATE: 18-NOV-1993
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/161,739
FILING DATE: 03-DEC-1993
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/165,699
FILING DATE: 10-DEC-1993
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/209,741
FILING DATE: 09-MAR-1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/352,322
FILING DATE: 07-DEC-1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/544,404
FILING DATE: 10-OCT-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/728,463
FILING DATE: 10-OCT-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/758,417
FILING DATE: 02-DEC-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US96/16433
FILING DATE: 10-OCT-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/21803
FILING DATE: 01-DEC-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 014643-009040US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-042-353-4

Query Match 100.0%; Score 7; DB 4; Length 12;

```


Best Local Similarity 100.0%; Pred. No. 9.1e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttgtata 7
|||||
Db 4 TTTGATA 10

RESULT 4

US-08-758-417A-272
; Sequence 272, Application US/08758417A
; Patent No. 6300129

GENERAL INFORMATION:

APPLICANT: Lomborg, Nils
Kay, Robert M.
TITLE OF INVENTION: Transgenic No. 6300129-Human Animals for
Producing Heterologous Antibodies

NUMBER OF SEQUENCES: 417

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/758,417A
FILING DATE: 02-Dec-1996

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/728,463
FILING DATE: 10-OCT-1996
APPLICATION NUMBER: US 08/544,404
FILING DATE: 10-OCT-1995
APPLICATION NUMBER: US 08/352,322
FILING DATE: 07-DEC-1994
APPLICATION NUMBER: US 08/209,741
FILING DATE: 09-MAR-1994
APPLICATION NUMBER: US 08/165,699
FILING DATE: 10-DEC-1993
APPLICATION NUMBER: US 08/161,739
FILING DATE: 03-DEC-1993
APPLICATION NUMBER: US 08/155,301
FILING DATE: 18-NOV-1993
APPLICATION NUMBER: US 08/096,762
FILING DATE: 22-JUL-1993
APPLICATION NUMBER: US 08/053,131
FILING DATE: 26-APR-1993
APPLICATION NUMBER: US 07/990,860
FILING DATE: 16-DEC-1992

ATTORNEY/AGENT INFORMATION:

NAME: Serafini, Andrew T.
REGISTRATION NUMBER: 41,303

REFERENCE/DOCKET NUMBER: 014643-00903005

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 272:

SEQUENCE CHARACTERISTICS:

LENGTH: 12 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

SEQUENCE DESCRIPTION: SEQ ID NO: 272:
US-08-758-417A-272

Query Match 100.0%; Score 7; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 9.1e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttgtata 7
|||||
Db 4 TTTGATA 10

RESULT 5

US-08-166-664-5/c
; Sequence 5, Application US/08166664
; Patent No. 5646020

GENERAL INFORMATION:

APPLICANT: James A. McSwiggen
Anthony Mamone
TITLE OF INVENTION: HAMMERHEAD RIBOZYMES FOR
PREFERRED TARGETS

NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Lyon & Lyon
STREET: 611 West Sixth Street
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90017

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/166,664
FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/07/884,074
FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 197/062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 13
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-166-664-5

Query Match 100.0%; Score 7; DB 1; Length 13;
Best Local Similarity 100.0%; Pred. No. 9.1e+03;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttgtata 7
|||||
Db 10 TTTGATA 4

RESULT 6

US-08-271-880A-217/c
; Sequence 217, Application US/08271880A
; Patent No. 5693535

GENERAL INFORMATION:

APPLICANT: Kenneth G. Draper
Bharat Chovvita
APPLICANT: James McSwiggen
APPLICANT: Dan T. Stinchcomb

APPLICANT: James D. Thompson
TITLE OF INVENTION: METHOD AND REAGENT FOR INHIBITING
TITLE OF INVENTION: HUMAN IMMUNODEFICIENCY VIRUS
TITLE OF INVENTION: REPLICATION
NUMBER OF SEQUENCES: 232
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/271,880A
FILING DATE: July 7, 1994
PRIOR APPLICATION DATA:
PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below:
APPLICATION NUMBER: 08/103,243
FILING DATE: August 6, 1993
APPLICATION NUMBER: 07/882,886
FILING DATE: May 14, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 206/116
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ. ID NO: 217:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-271-880A-217

Query Match 100.0%; Score 7; DB 1; Length 13;
Best Local Similarity 100.0%; Pred. No. 9.1e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ttgata 7
1111111
DB 10 TTTGATA 4

RESULT 7
US-08-910-408-217/C
Sequence 217, Application US/08910408
Patent No. 5972704
GENERAL INFORMATION:
APPLICANT: Kenneth G. Draper
ADDRESSEE: Bharat Chowrira
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/271,880A
FILING DATE: July 7, 1994
PRIOR APPLICATION DATA:
PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below:
APPLICATION NUMBER: 08/103,243
FILING DATE: August 6, 1993
APPLICATION NUMBER: 07/882,886
FILING DATE: May 14, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 206/116
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ. ID NO: 217:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-910-408-217

CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/910,408
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/271,880
FILING DATE: July 7, 1994
APPLICATION NUMBER: 08/103,243
FILING DATE: August 6, 1993
APPLICATION NUMBER: 07/882,886
FILING DATE: May 14, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 206/116
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ. ID NO: 217:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-910-408-217

Query Match 100.0%; Score 7; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 9.1e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ttgata 7
1111111
DB 10 TTTGATA 4

RESULT 8
US-09-249-215-217/C
Sequence 217, Application US/09249215
Patent No. 6159692
GENERAL INFORMATION:
APPLICANT: Kenneth G. Draper
ADDRESSEE: Bharat Chowrira
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/910,408
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/271,880
FILING DATE: July 7, 1994
APPLICATION NUMBER: 08/103,243
FILING DATE: August 6, 1993
APPLICATION NUMBER: 07/882,886
FILING DATE: May 14, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 206/116
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ. ID NO: 217:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-910-408-217

SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/249,215
FILING DATE: 12-Feb-1999
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/910,408
FILING DATE: <Unknown>
APPLICATION NUMBER: 08/103,243
FILING DATE: August 6, 1993
APPLICATION NUMBER: 07/882,886
FILING DATE: May 14, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 206/116
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 217:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 217:
US-09-249-215-217

Query Match 100.0%; Score 7; DB 3; Length 13;
Best Local Similarity 100.0%; Pred. No. 9.1e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ttgata 7
|||||
Db 10 TTGATA 4

RESULT 9
US-08-271-880A-212
Sequence 212, Application US/08271880A
Patent No. 5693535
GENERAL INFORMATION:
APPLICANT: Kenneth G. Draper
APPLICANT: Bharat Chowrira
APPLICANT: James McSwiggen
APPLICANT: Dan T. Stinchcomb
APPLICANT: James D. Thompson
TITLE OF INVENTION: METHOD AND REAGENT FOR INHIBITING
TITLE OF INVENTION: HUMAN IMMUNODEFICIENCY VIRUS
NUMBER OF SEQUENCES: 232
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/271,880A
FILING DATE: July 7, 1994
PRIOR APPLICATION DATA:
PRIOR APPLICATION DATA: including application
APPLICATION NUMBER: 08/103,243

FILING DATE: August 6, 1993
APPLICATION NUMBER: 07/882,886
FILING DATE: May 14, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 206/116
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 212:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-271-880A-212

Query Match 100.0%; Score 7; DB 1; Length 14;
Best Local Similarity 100.0%; Pred. No. 9.1e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ttgata 7
|||||
Db 3 TTGATA 9

RESULT 10
US-08-910-408-212
Sequence 212, Application US/08910408
Patent No. 5972704
GENERAL INFORMATION:
APPLICANT: Kenneth G. Draper
APPLICANT: Bharat Chowrira
APPLICANT: James McSwiggen
APPLICANT: Dan T. Stinchcomb
APPLICANT: James D. Thompson
TITLE OF INVENTION: METHOD AND REAGENT FOR INHIBITING
TITLE OF INVENTION: HUMAN IMMUNODEFICIENCY VIRUS
NUMBER OF SEQUENCES: 232
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/910,408
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/271,880
FILING DATE: July 7, 1994
APPLICATION NUMBER: 08/103,243
FILING DATE: August 6, 1993
APPLICATION NUMBER: 07/882,886
FILING DATE: May 14, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 206/116
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600

TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 212:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-910-408-212

Query Match 100.0%; Score 7; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 9.1e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ttgata 7
|||||||
DB 3 TTGATA 9

RESULT 11
US-09-249-215-212
Sequence 212, Application US/09249215
Patent No. 6159692
GENERAL INFORMATION:
APPLICANT: Kenneth G. Draper
Bharat Chowilla
James McSwiggen
Dan T. Stinchcomb
James D. Thompson
TITLE OF INVENTION: METHOD AND REAGENT FOR INHIBITING
HUMAN IMMUNODEFICIENCY VIRUS
REPLICATION
NUMBER OF SEQUENCES: 232
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/249,215
FILING DATE: 12-Feb-1999
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/910,408
FILING DATE: <unknown>
APPLICATION NUMBER: 08/103,243
FILING DATE: August 6, 1993
APPLICATION NUMBER: 07/882,886
FILING DATE: May 14, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 206/116
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 212:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 212:

US-09-249-215-212

Query Match 100.0%; Score 7; DB 3; Length 14;
Best Local Similarity 100.0%; Pred. No. 9.1e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ttgata 7
|||||||
DB 3 TTGATA 9

RESULT 12
US-08-373-124A-90/c
Sequence 90, Application US/08373124A
Patent No. 5646042
GENERAL INFORMATION:
APPLICANT: Stinchcomb, Dan T.
APPLICANT: Draper, Kenneth
APPLICANT: McSwiggen, James
APPLICANT: Jarvis, Thale
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
TREATMENT OF RESPONOSIS AND
TITLE OF INVENTION: CANCER USING RIBOZYMES
NUMBER OF SEQUENCES: 2627
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/373,124A
FILING DATE: January 13, 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/245,466
FILING DATE: May 18, 1994
APPLICATION NUMBER: 08/192,943
FILING DATE: February 7, 1994
APPLICATION NUMBER: 07/987,132
FILING DATE: December 7, 1992
APPLICATION NUMBER: 07/936,422
FILING DATE: August 26, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 209/035
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 90:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-373-124A-90

Query Match 100.0%; Score 7; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 9.1e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ttgata 7

Db 11 TTTGATA 5

RESULT 13

US-08-291-932A-129
Sequence 129, Application US/08291932A
Patent No. 5658780

GENERAL INFORMATION:

APPLICANT: Stinchcomb, Dan T.
APPLICANT: Draper, Kenneth G.
APPLICANT: McSwigen, James
TITLE OF INVENTION: RIBOZYME TREATMENT OF
DISEASES OR CONDITIONS
TITLE OF INVENTION: RELATED TO LEVELS OF
TITLE OF INVENTION: NF-KB
NUMBER OF SEQUENCES: 830
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/291,932A
FILING DATE: August 15, 1994

CLASSIFICATION: 514

PRIOR APPLICATION DATA: Including Application
PRIOR APPLICATION DATA: described below:
APPLICATION NUMBER: 08/245,466
FILING DATE: May 18, 1994

Two

APPLICATION NUMBER: 07/987,132
FILING DATE: December 7, 1992

ATTORNEY/AGENT INFORMATION:

NAME: Wardburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 208/157
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440

TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 129:

SEQUENCE CHARACTERISTICS:

LENGTH: 15 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-291-932A-129

Query Match

Best Local Similarity 100.0%; Score 7; DB 1; Length 15;
Matches 3; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Oy 1 tttagata 7

Db 9 UUGCAUA 15

RESULT 14

US-08-291-932A-130
Sequence 130, Application US/08291932A
Patent No. 5658780
GENERAL INFORMATION:

APPLICANT: Stinchcomb, Dan T.
APPLICANT: Draper, Kenneth G.
APPLICANT: McSwigen, James

TITLE OF INVENTION: RIBOZYME TREATMENT OF
DISEASES OR CONDITIONS
TITLE OF INVENTION: RELATED TO LEVELS OF
TITLE OF INVENTION: NF-KB
NUMBER OF SEQUENCES: 830
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/291,932A
FILING DATE: August 15, 1994

CLASSIFICATION: 514

PRIOR APPLICATION DATA: Including application
PRIOR APPLICATION DATA: described below:
APPLICATION NUMBER: 08/245,466
FILING DATE: May 18, 1994

Two

APPLICATION NUMBER: 07/987,132
FILING DATE: December 7, 1992

ATTORNEY/AGENT INFORMATION:

NAME: Wardburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 208/157
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440

TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 130:

SEQUENCE CHARACTERISTICS:
LENGTH: 15 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-291-932A-130

Query Match

Best Local Similarity 100.0%; Score 7; DB 1; Length 15;
Matches 3; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Oy 1 tttagata 7

Db 7 UUGCAUA 13

RESULT 15

US-08-291-932A-131
Sequence 131, Application US/08291932A
Patent No. 5658780

GENERAL INFORMATION:

APPLICANT: Stinchcomb, Dan T.
APPLICANT: Draper, Kenneth G.
APPLICANT: McSwigen, James
TITLE OF INVENTION: RIBOZYME TREATMENT OF
DISEASES OR CONDITIONS
TITLE OF INVENTION: RELATED TO LEVELS OF
TITLE OF INVENTION: NF-KB
NUMBER OF SEQUENCES: 830
CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon
 STREET: 633 West Fifth Street
 STREET: Suite 4700
 CITY: Los Angeles
 STATE: California
 COUNTRY: U.S.A.
 ZIP: 90071-2066
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 MEDIUM TYPE: storage
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: IBM P.C. DOS 5.0
 SOFTWARE: Word Perfect 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/291,932A
 FILING DATE: August 15, 1994
 CLASSIFICATION: 514
 PRIOR APPLICATION DATA:
 PRIOR APPLICATION DATA: including application
 PRIOR APPLICATION DATA: described below:
 APPLICATION NUMBER: 08/245,466
 FILING DATE: May 18, 1994
 APPLICATION NUMBER: 07/987,132
 FILING DATE: December 7, 1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Warburg, Richard J.
 REGISTRATION NUMBER: 32,327
 REFERENCE/DOCKET NUMBER: 208/157
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (213) 489-1600
 TELEFAX: (213) 955-0440
 TELEEX: 67-3510
 INFORMATION FOR SEQ ID NO: 131:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 15 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-08-291-932A-131

Two

Query Match 100.0%; Score 7; DB 1; Length 15;
 Best Local Similarity 42.9%; Pred. No. 9.1e+03;
 Matches 3; Conservative 4; Mismatches 0; Indels 0;
 Gaps 0;
 OY 1 ttgata 7
 Db 6 UUGAUA 12

Search completed: December 15, 2001, 03:22:47
 Job time: 7562 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 02:33:41 : Search time 4441.54 Seconds
(without alignments)
16.936 Million cell updates/sec

Title: US-09-380-826a-5

Perfect score: 7

Sequence: 1 ttgata 7

Scoring table: IDENTITY_NUC

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

EST: *
1: em_estfun: *
2: em_esthum: *
3: em_estln: *
4: em_estom: *
5: em_estpl: *
6: em_estba: *
7: em_estro: *
8: em_estov: *
9: em_hlc: *
10: gb_est1: *
11: gb_est2: *
12: gb_hlc: *
13: gb_gss: *
14: em_gss_fun: *
15: em_gss_hum: *
16: em_gss_inv: *
17: em_gss_pln: *
18: em_gss_pro: *
19: em_gss_rtd: *
20: em_gss_vrt: *
21: em_gss_other: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
C 1	7	100.0	19	AZ641145	AZ641145 IM0503F04
C 2	7	100.0	20	AZ323820	AZ323820 IM0045A01
C 3	7	100.0	20	AZ819886	AZ819886 2M0091H16
C 4	7	100.0	20	AZ977536	AZ977536 2M0253B04
C 5	7	100.0	22	AZ329387	AZ329387 IM0053C07
C 6	7	100.0	22	AZ502952	AZ502952 IM0342C04
C 7	7	100.0	22	AZ828775	AZ828775 2M0106B06
C 8	7	100.0	23	AZ345476	AZ345476 IM0080P11
C 9	7	100.0	23	AZ789315	AZ789315 2M0036M22
C 10	7	100.0	23	AZ840208	AZ840208 2M0136H12
C 11	7	100.0	23	AZ862305	AZ862305 2M0165N22
C 12	7	100.0	23	TA144D04Q	AL466870 T. brucei

C 13	7	100.0	24	AZ379215	AZ379215 IM0134M14
C 14	7	100.0	24	AZ480650	AZ480650 IM0302P17
C 15	7	100.0	24	AZ663958	AZ663958 IM0543R19
C 16	7	100.0	24	AZ794696	AZ794696 2M0048C04
C 17	7	100.0	25	AU008929	AU008929 AU0089929
C 18	7	100.0	25	AZ404619	AZ404619 IM0173L20
C 19	7	100.0	25	AZ592455	AZ592455 IM0403H19
C 20	7	100.0	26	AW333097	AW333097 S17C3 AGS
C 21	7	100.0	27	D18735	D18735 MUSGSS01797
C 22	7	100.0	27	AZ328549	AZ328549 IM0052D24
C 23	7	100.0	27	AZ335603	AZ335603 IM0065E13
C 24	7	100.0	27	AZ836137	AZ836137 2M0130M13
C 25	7	100.0	27	AZ966573	AZ966573 2M0237B12
C 26	7	100.0	28	AZ850027	AZ850027 2M0151P09
C 27	7	100.0	29	AZ966795	AZ966795 2M0237P23
C 28	7	100.0	29	TA97B05P	AL461010 T. brucei
C 29	7	100.0	30	AU007442	AU007442 AU007442
C 30	7	100.0	30	C20899	C20899 HUMGS000497
C 31	7	100.0	30	BE911076	BE911076 601662175
C 32	7	100.0	30	AZ601864	AZ601864 IM0420P16
C 33	7	100.0	30	TA118C09P	AL463710 T. brucei
C 34	7	100.0	31	A1140482	A1140482 Cw82d12.S
C 35	7	100.0	31	A1973656	A1973656 sc07h11.Y
C 36	7	100.0	31	AZ591446	AZ591446 IM0401J05
C 37	7	100.0	31	AZ799349	AZ799349 2M0056G12
C 38	7	100.0	32	AU009849	AU009849 AU009849
C 39	7	100.0	32	D18230	D18230 MUSGSS00509
C 40	7	100.0	32	AZ447488	AZ447488 IM0244K07
C 41	7	100.0	33	R98634	R98634 Yq63H06.r1
C 42	7	100.0	33	AZ308864	AZ308864 IM0012M13
C 43	7	100.0	33	AZ411435	AZ411435 IM0184C05
C 44	7	100.0	33	AZ768977	AZ768977 IM0569B14
C 45	7	100.0	34	TA164G09Q	AL473172 T. brucei

ALIGNMENTS

RESULT 1
LOCUS AZ641145/c 19 bp DNA
DEFINITION IM0503F04R Mouse 10kb plasmid UUGCIM library Mus musculus genomic
clone UUGCIM0503F04 R, DNA sequence.
ACCESSION AZ641145
VERSION AZ641145.1 GI:11764838
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 19)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamli,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tingley,A., von Niederhausern,A.
and Wright,D., Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0503 row: F column: 04
Seq primer: CACACAGCAACACGCTATGACC
Class: plasmid ends
High quality sequence stop: 19.
Location/Qualifiers
1. 19

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0503F04"
/clone.lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g1473211419b1AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT
ORIGIN
7 a 5 c 2 g 5 t

Query Match 100.0%; Score 7; DB 13; Length 19;
Best local Similarity 100.0%; Pred. No. 4.2e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ttgtgata 7
|||||||
Db 19 TTTGATTA 13

RESULT 2
A2323820/c 20 bp DNA GSS 29-SEP-2000
LOCUS 1M0045A01R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
DEFINITION clone UUGC1M0045A01 R, DNA sequence.
ACCESSION A2323820
VERSION A2323820.1 GI:10378917
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 20)
REFERENCE
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0045 row: A column: 01
Seq primer: CACACAGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 20.
Location/Qualifiers

Source
1. 20
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0045A01"
/clone.lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g1473211419b1AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT
ORIGIN
10 a 4 c 3 g 3 t

Query Match 100.0%; Score 7; DB 13; Length 20;
Best local Similarity 100.0%; Pred. No. 4.2e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ttgtgata 7
|||||||
Db 11 TTTGATA 5

RESULT 3
A2819886 20 bp DNA GSS 20-FEB-2001
LOCUS 2M0091H16R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
DEFINITION clone UUGC2M0091H16 R, DNA sequence.
ACCESSION A2819886
VERSION A2819886.1 GI:12989794
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 20)
REFERENCE
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0091 row: H column: 16
Seq primer: CACACAGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 20.

FEATURES
source
Location/Qualifiers
1. .20
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="U06C2M0253B04"
/clone_lib="Mouse 10kb plasmid U06C2M library"
/sex="Male"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (g114732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT
ORIGIN
7 a 3 c 4 g 6 t

Query Match
Best Local Similarity 100.0%; Score 7; DB 13; Length 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ttgtgata 7
|||||||
Db 10 TTTGATA 16

RESULT 4
A2977536 20 bp DNA GSS 27-APR-2001
LOCUS 2M0253B04R Mouse 10kb plasmid U06C2M library Mus musculus genomic
DEFINITION clone U06C2M0253B04 R, DNA sequence.
ACCESSION A2977536
VERSION A2977536.1 GI:13848763
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 20)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Roge,R., Stokes,R., Tinney,A., von Niederhausern,A. and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert length: 10000 Std Error: 0.00
Plate: 0253 row: B column: 04
Seq primer: CACACAGGAACACGTATGACC
Class: plasmid ends

JOURNAL
COMMENT

High quality sequence stop: 20.
FEATURES
source
Location/Qualifiers
1. .20
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="U06C2M0253B04"
/clone_lib="Mouse 10kb plasmid U06C2M library"
/sex="Female"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (female) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (g114732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT
ORIGIN
4 a 5 c 2 g 9 t

Query Match
Best Local Similarity 100.0%; Score 7; DB 13; Length 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ttgtgata 7
|||||||
Db 7 TTTGATA 13

RESULT 5
A2329387 22 bp DNA GSS 29-SEP-2000
LOCUS 1M0053C07R Mouse 10kb plasmid U06C1M library Mus musculus genomic
DEFINITION clone U06C1M0053C07 R, DNA sequence.
ACCESSION A2329387
VERSION A2329387.1 GI:10390050
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 22)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Roge,R., Stokes,R., Tinney,A., von Niederhausern,A. and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert length: 10000 Std Error: 0.00
Plate: 0053 row: C column: 07
Seq primer: CACACAGGAACACGTATGACC

JOURNAL
COMMENT

Class: plasmid ends
High quality sequence stop: 22.
Location/Qualifiers
source

1. 22
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0053C07"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42ny; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g11473211419b1AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 8 a 4 c 3 g 7 t
ORIGIN

Query Match 100.0%; Score 7; DB 13; Length 22;
Best Local Similarity 100.0%; Pred. No. 4.2e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ttgtgata 7
|||||||
Db 21 TTTGATA 15

RESULT 6
AZ502952 22 bp DNA GSS 05-OCT-2000
LOCUS 1M0342C04R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
DEFINITION clone UUGC1M0342C04 R, DNA sequence.
ACCESSION AZ502952
VERSION AZ502952.1 GI:10684268
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 22)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
Unpublished (2000)
JOURNAL Contact: Robert B. Weiss
COMMENT University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0342 row: C column: 04

Seq primer: CACACAGAAACGCTATGACC
Class: plasmid ends
High quality sequence stop: 22.
Location/Qualifiers
source

1. 22
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0342C04"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42ny; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g11473211419b1AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 10 a 5 c 1 g 6 t
ORIGIN

Query Match 100.0%; Score 7; DB 13; Length 22;
Best Local Similarity 100.0%; Pred. No. 4.2e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ttgtgata 7
|||||||
Db 9 TTTGATA 3

RESULT 7
AZ828775 22 bp DNA GSS 20-FEB-2001
LOCUS 2M0106B06F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
DEFINITION clone UUGC2M0106B06 F, DNA sequence.
ACCESSION AZ828775
VERSION AZ828775.1 GI:12998683
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 22)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
Unpublished (2000)
JOURNAL Contact: Robert B. Weiss
COMMENT University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00

Plate: 0106 row: B column: 06
Seq primer: CGTGTAAACGACGCCACGT
Class: plasmid ends
High quality sequence stop: 22.

FEATURES
source

1. 22
Location/Qualifiers
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="U06C2M0106B06"
/clone_lib="Mouse 10Kb plasmid U06C1M library"
/sex="Male"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g114732114(gblAF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT
ORIGIN

9 a 7 c 2 g 4 t

Query Match 100.0%; Score 7; DB 13; Length 22;
Best Local Similarity 100.0%; Pred. No. 4.2e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ttgata 7
|||||

Db 21 TTTGATA 15

RESULT 8
A2345476/c

LOCUS

DEFINITION

A2345476 23 bp DNA GSS 29-SEP-2000
1M0080P11F Mouse 10kb plasmid U06C1M library Mus musculus genomic
clone U06C1M0080P11 F, DNA sequence.

ACCESSION

A2345476
A2345476.1 GI:10424713

VERSION

KEYWORDS

SOURCE

ORGANISM

house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 23)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10Kb
plasmid inserts

TITLE

JOURNAL

COMMENT

Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00
Plate: 0080 row: P column: 11
Seq primer: CGTGTAAACGACGCCACGT
Class: plasmid ends
High quality sequence stop: 23.

FEATURES
source

1. 23
Location/Qualifiers
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="U06C1M0080P11"
/clone_lib="Mouse 10Kb plasmid U06C1M library"
/sex="Male"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g114732114(gblAF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT
ORIGIN

12 a 1 c 0 g 10 t

Query Match 100.0%; Score 7; DB 13; Length 23;
Best Local Similarity 100.0%; Pred. No. 4.2e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ttgata 7
|||||

Db 23 TTTGATA 17

RESULT 9
A2789315/c

LOCUS

DEFINITION

A2789315 23 bp DNA GSS 16-FEB-2001
2M0036W2R Mouse 10kb plasmid U06C1M library Mus musculus genomic
clone U06C2M0036W22 R, DNA sequence.

ACCESSION

A2789315
A2789315.1 GI:12929986

VERSION

KEYWORDS

SOURCE

house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 23)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10Kb
plasmid inserts

TITLE

JOURNAL

COMMENT

Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177

Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0036 row: M column: 22
 Seq primer: CACACAGAAACAGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 23.
 Location/Qualifiers

FEATURES

source

1. 23
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC2M036M22"
 /clone_1lb="Mouse 10kb plasmid UUGC1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g11473211419b/AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT

9 a 4 c 4 g 6 t

ORIGIN

Query Match 100.0%; Score 7; DB 13; Length 23;
 Best Local Similarity 100.0%; Pred. No. 4.2e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ttgtgata 7
 |||||
 Db 22 TTGTGATA 16

RESULT 10
 A2840208 23 bp DNA GSS 20-FEB-2001
 LOCUS A2840208/c
 DEFINITION 2M0136H12R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
 accession A2840208
 version A2840208.1 GI:13010116
 keywords GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 23)
 Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C.,
 Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
 M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
 and Wright,D., Weiss,R.
 Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 Unpublished (2000)

JOURNAL COMMENT
 Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606

Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0136 row: H column: 12
 Seq primer: CACACAGAAACAGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 23.
 Location/Qualifiers

FEATURES

source

1. 23
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC2M0136H12"
 /clone_1lb="Mouse 10kb plasmid UUGC1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g11473211419b/AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT

8 a 7 c 3 g 5 t

ORIGIN

Query Match 100.0%; Score 7; DB 13; Length 23;
 Best Local Similarity 100.0%; Pred. No. 4.2e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ttgtgata 7
 |||||
 Db 17 TTGTGATA 11

RESULT 11
 A2862305 23 bp DNA GSS 21-FEB-2001
 LOCUS A2862305/c
 DEFINITION 2M0169N22F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
 accession A2862305
 version A2862305.1 GI:13059475
 keywords GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 23)
 Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C.,
 Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
 M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
 and Wright,D., Weiss,R.
 Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 Unpublished (2000)

JOURNAL COMMENT
 Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA

Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0169 row: N column: 22
 Seq primer: CGTTGTAACGACGCCACGT
 Class: plasmid ends
 High quality sequence stop: 23.
 Location/Qualifiers

FEATURES

SOURCE

1. 23
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUCG2M0169N22"
 /clone_1lb="Mouse 10kb plasmid UUCG1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (g11473211419b1AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT

7 a 7 c 1 g 8 t

ORIGIN

Query Match 100.0%; Score 7; DB 13; Length 23;
 Best Local Similarity 100.0%; Pred. No. 4.2e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 ttgtgata 7
 |||||
 Db 17 TTTGATA 11

RESULT 12
 TAI14D040 23 bp DNA GSS 13-DEC-2000
 LOCUS T. brucei sheared genomic DNA clone 144d04, reverse sequence,
 DEFINITION genomic survey sequence.
 ACCESSION AL466870
 VERSION AL466870.1 GI:11836225
 KEYWORDS GSS.
 SOURCE Trypanosoma brucei.
 ORGANISM Trypanosoma brucei
 Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
 Trypanosoma.

REFERENCE

AUTHORS

1 (bases 1 to 23)
 Hall, N., Bowman, S., Lennard, N.J., Doggett, J., Atkin, R.,
 Chillingworth, C., Ormond, D., Harris, B., El-Sayed, N., Hou, L.,
 Melville, S.E., Rajandream, M.A. and Barrell, B.G.

TITLE

JOURNAL

Direct Submission
 Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
 project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
 Cambridge CB10 1SA, Email: barrell@sanger.ac.uk and
 nh@sanger.ac.uk

COMMENT

Constructed at the Institute for Genomic Research (TIGR),
 Rockville, MD. Genomic DNA isolated from a cloned population of
 Trypanosoma brucei (TREU927/4 G9rat 10.1) was mechanically sheared

to give a tight size distribution (4 kb). The v + i method used for the library construction is described in detail in Smith, H. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Barrell, Oxford University Press, 1999).
 Email: nelsayed@tigr.org
 Details of T. brucei sequencing at the Sanger Centre are available at <http://www.sanger.ac.uk/projects/T-brucei/>.

FEATURES

SOURCE

1. 23
 /organism="Trypanosoma brucei"
 /strain="TREU927"
 /db_xref="taxon:5691"
 /clone="144d04"

BASE COUNT

3 a 6 c 5 g 9 t

ORIGIN

Query Match 100.0%; Score 7; DB 13; Length 23;
 Best Local Similarity 100.0%; Pred. No. 4.2e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 ttgtgata 7
 |||||
 Db 11 TTTGATA 17

RESULT 13
 AZ379215/c 24 bp DNA GSS 02-OCT-2000
 LOCUS M0134M14F Mouse 10kb plasmid UUCG1M library Mus musculus genomic
 DEFINITION clone UUCG1M0134M14 F, DNA sequence.
 ACCESSION AZ379215
 VERSION AZ379215.1 GI:10492915
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 24)
 Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C.,
 Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
 M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.
 and Wright, D., Weiss, R.
 Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 JOURNAL Unpublished (2000)
 CONTACT: Robert B. Weiss
 UNIVERSITY OF UTAH
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0134 row: M column: 14
 Seq primer: CGTTGTAACGACGCCACGT
 Class: plasmid ends
 High quality sequence stop: 24.
 Location/Qualifiers

FEATURES

SOURCE

1. 24
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUCG1M0134M14"
 /clone_1lb="Mouse 10kb plasmid UUCG1M library"

/sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g11473211419b|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 8 a 4 c 4 g 8 t
ORIGIN

Query Match 100.0%; Score 7; DB 13; Length 24;
Best Local Similarity 100.0%; Pred. No. 4.2e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttgata 7
111111
Db 12 TTTGATA 6

RESULT 14
AZ480650 24 bp DNA GSS 04-OCT-2000
LOCUS 1M0302P17F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
DEFINITION clone UUGC1M0302P17 F, DNA sequence.
ACCESSION AZ480650
VERSION AZ480650.1 GI:10641715
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert length: 10000 Std Error: 0.00
Plate: 0302 row: P column: 17
Seq primer: CGTGTAAACGACGCGCAGT
Class: plasmid ends
High quality sequence stop: 24.
Location/Qualifiers

FEATURES
source 1..24

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0302P17"
/clone_11b="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson

Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g11473211419b|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 11 a 4 c 1 g 8 t
ORIGIN

Query Match 100.0%; Score 7; DB 13; Length 24;
Best Local Similarity 100.0%; Pred. No. 4.2e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttgata 7
111111
Db 8 TTTGATA 2

RESULT 15
AZ663958 24 bp DNA GSS 14-DEC-2000
LOCUS 1M0543E19R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
DEFINITION clone UUGC1M0543E19 R, DNA sequence.
ACCESSION AZ663958
VERSION AZ663958.1 GI:11801104
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert length: 10000 Std Error: 0.00
Plate: 0543 row: E column: 19
Seq primer: CACACAGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 24.
Location/Qualifiers

FEATURES
source 1..24

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0543E19"
/clone_11b="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g11473214|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 11 a 4 c 2 g 7 t
ORIGIN

Query Match 100.0%; Score 7; DB 13; Length 24;
Best Local Similarity 100.0%; Pred. No. 4.2e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 tttagata 7
|||||||
Db 24 TTTGATA 18

Search completed: December 15, 2001, 02:33:44
Job time: 4984 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 03:19:53 : Search time 2725.73 Seconds
(without alignments)
133.153 Million cell updates/sec

Title: US-09-380-826a-6

Perfect score: 1 tgttgannnnnnnttgata 22

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*
1: gb_da:*
2: gb_htg:*
3: gb_in:*
4: gb_ov:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_ro:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_ov:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vl:*
30: em_htgo_hum:*
31: em_htgo_inv:*
32: em_htg_hum:*
33: em_htg_hum:*
34: em_htg_inv:*
35: em_htg_inv:*
36: em_htg_other:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	14	63.6	239	9	HSSSG03	AF051649 Homo sapi
2	14	63.6	435	11	HUM4STS161	M95263 Human Chrom
3	14	63.6	668	11	G50458	G50458 SHC-79916
4	14	63.6	1015	14	SRSSRSV	D38547 Small round
5	14	63.6	1257	4	BTAG9225	AJ009225 Bos tauru
6	14	63.6	1257	9	HSAG9229	AJ009229 Homo sapi
7	14	63.6	1258	4	BTAG9224	AJ009224 Bos tauru
8	14	63.6	1276	4	BTAG9223	AJ009223 Bos tauru
9	14	63.6	1287	9	HSUS2077	U52077 Human marin
10	14	63.6	1305	6	AXI48809	AXI48809 Sequence
11	14	63.6	1306	6	AXI48806	AXI48806 Sequence
12	14	63.6	1481	1	LFU60594	U60594 Leptosira
13	14	63.6	2321	3	DDU66913	U66913 Dictyostell
14	14	63.6	2683	1	SPGROELN	X89236 S.pyogenes
15	14	63.6	2683	6	AXI48805	AXI48805 Sequence
16	14	63.6	3666	6	A45819	A45819 Sequence 13
17	14	63.6	3666	6	A45820	A45820 Sequence 14
18	14	63.6	3915	8	SCYJL046W	Z49321 S.cerevisia
19	14	63.6	4197	6	A45813	A45813 Sequence 7
20	14	63.6	4197	6	A45814	A45814 Sequence 8
21	14	63.6	4684	8	SCYJL045W	Z49320 S.cerevisia
22	14	63.6	5620	10	MTTHREC02	U36757 Mus musculu
23	14	63.6	7697	14	AB042808	AB042808 Chiba vir
24	14	63.6	10389	1	AE006627	AE006627 Streptoco
25	14	63.6	10948	2	AC014767	AC014767 Drosophi
26	14	63.6	11949	9	AF000145	AF000145 Homo sapi
27	14	63.6	12437	1	AE002102	AE002102 Ureaplas
28	14	63.6	12973	2	AC013026	AC013026 Drosophi
29	14	63.6	12978	3	AC006805	AC006805 Caenorhab
30	14	63.6	26923	3	CEM0386	Z78545 Caenorhabd
31	14	63.6	29000	3	AY027893S6	AY027898 Homo sapi
32	14	63.6	37923	3	AF047659	AF047659 Caenorhab
33	14	63.6	42042	3	CEC3484	Z78059 Caenorhabd
34	14	63.6	59012	9	HSABLGR2	U07562 Human Abl.9
35	14	63.6	59909	2	AC090996	AC090996 Homo sapi
36	14	63.6	63156	2	AC026482	AC026482 Homo sapi
37	14	63.6	63659	2	AC090590	AC090590 Homo sapi
38	14	63.6	63739	9	AF271897	AF271897 Homo sapi
39	14	63.6	63933	2	AC090570	AC090570 Homo sapi
40	14	63.6	64577	3	AE002828	AE002828 Drosophi
41	14	63.6	64789	2	AC083839	AC083839 Homo sapi
42	14	63.6	67987	2	AC016064	AC016064 Homo sapi
43	14	63.6	68082	2	AC014851	AC014851 Drosophi
44	14	63.6	68206	2	AC087450	AC087450 Homo sapi
45	14	63.6	69358	2	AC090544	AC090544 Homo sapi

ALIGNMENTS

RESULT 1
LOCUS HSSSG03 239 bp DNA PRI 14-MAR-1999
DEFINITION Homo sapiens squalene synthase gene, intron 2, 5' end.
ACCESSION AF051649
VERSION AF051649.1 GI:4415976

KEYWORDS
SEGMENTS
SOURCE
ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 239)
Jiang, G., McKenzie, T.L., Conrad, D.G. and Shechter, I.

Transcriptional regulation by lovastatin and 25-hydroxycholesterol
in HepG2 cells and molecular cloning and expression of the cDNA for
the human hepatic squalene synthase

JOURNAL J. Biol. Chem. 268 (17), 12818-12824 (1993)

MEDLINE 93286128
REFERENCE 2 (bases 1 to 239)

AUTHORS Guan, G., Dai, P.H. and Shechter, I.
 TITLE Direct Submission
 JOURNAL Submitted (02-MAR-1998) Biochemistry and Molecular Biology, Uniformed Services University of the Health Sciences, 4301 Jones Bridge Road, Bethesda, MD 20814-4799, USA

FEATURES
 source 1. 239
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="8"
 /map="8p22-p23.1"
 33. .>239
 /gene="squalene synthase"
 /number=2
 Intron

BASE COUNT 43 a 61 c 70 g 65 t

ORIGIN

Query Match 63.6%; Score 14; DB 9; Length 239;
 Best Local Similarity 63.6%; Pred. No. 1.2e+03;
 Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Oy 1 tgttgagannnnnnnttgata 22
 ||||| |||||
 Db 92 TGTTCGAGCTACTCTTGATA 113

RESULT 2
 HUM4STS161 435 bp DNA STS 26-FEB-1996
 LOCUS Hum. chromosome 4 sequence-tagged site STS4-161, sequence tagged
 DEFINITION site.
 ACCESSION M95263.1 GI:177262
 VERSION
 KEYWORDS STS; human chromosome 4; sequence tagged site.
 SOURCE Homo sapiens, clone C4-177 from Los Alamos National Laboratory chromosome 4 cosmid library, plate 4-1-2R, pos3c; vector SuperCos 1 (Strataene).

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 435)
 Goold, R.D., disibio, G., Xu, H., Lang, D.B., Dadgar, J., Magrane, G., Dugalczyk, A., Smith, K.A., Cox, D.R., Masters, S.B. and Myers, R.M.
 The development of sequence-tagged sites for human chromosome 4 Hum. Mol. Genet. 2 (8), 1271-1288 (1993)
 94004872

COMMENT
 PCR components: 25 ng of human genomic DNA, 10 pmol of each oligonucleotide, 200 micro-M dNTPs, 0.25 U Taq polymerase (cetus) in 10 micro-1 of 50 mM KCl-20 mM Tris-HCl, pH 8.3 (at room temp), 2.5 mM MgCl-2. Initial denaturation at 94degC for 1.5 min, then 30 cycles of 94degC for 15 sec, 62degC for 23 sec, and 72degC for 30 sec, followed by a final extension at 72degC for 3.5 min, using a Perkin-Elmer 9600 thermocycler. PCR-amplified product size 279 bp.
 Sequence submitted by:
 Human Genome Mapping Center
 Box 0925
 University of California San Francisco
 San Francisco, CA 94143-0925 USA
 Phone: (415) 502-1612 Fax: (415) 476-8391
 e-mail: hgmprobe@cgl.ucsf.edu.

FEATURES
 source 1. 435
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 primer_bind 316. .340
 primer_bind complement(62. .84)
 BASE COUNT 128 a 64 c 75 g 162 t 6 others
 ORIGIN

Query Match 63.6%; Score 14; DB 11; Length 435;
 Best Local Similarity 63.6%; Pred. No. 1.2e+03;

Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Oy 1 tgttgannnnnnnttgata 22
 ||||| |||||
 Db 262 TGTTCGAAATTGCTGTGATA 283

RESULT 3
 G50458 668 bp DNA STS 30-MAR-2000
 LOCUS SHGC-79916 Human Homo sapiens STS genomic, sequence tagged site.
 DEFINITION G50458
 ACCESSION G50458.1 GI:5221635
 VERSION
 KEYWORDS STS.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 668)
 Olivier, M. and Cox, D.R.
 Unpublished, Olivier, M., Cox, D.R. (2000)
 Unpublished (2000)

REFERENCE
 AUTHORS Olivier, M. and Cox, D.R.
 TITLE Unpublished, Olivier, M., Cox, D.R. (2000)
 JOURNAL Unpublished (2000)
 COMMENT

Contact: Michael Olivier, David R. Cox
 Stanford Human Genome Center
 Stanford University, School of Medicine
 4005 Miranda Ave. 2nd Fl., Palo Alto, CA 94025, USA
 Tel: (650) 320-5800
 Fax: (650) 320-5801
 Email: olivier@shgc.stanford.edu
 Primer A: GCACATTAGCTTGAGATGCTT
 Primer B: TGCCTTGTGCTTACCAATTTT
 STS size: 278
 PCR Profile:
 Initial incubation: 95 degrees C for 10 minutes
 Denaturation: 94 degrees C for 30 seconds
 Annealing: 60 degrees C for 30 seconds
 Polymerization: 72 degrees C for 23 seconds
 PCR Cycles: 30
 Thermal Cycler: Perkin Elmer 9700
 Protocol:
 Template: 25 ng
 Primer: each 1 uM
 dNTPs: each 200 uM
 AmpliTaq Gold Polymerase: 0.07 units/uL
 Total Vol: 5 uL
 Buffer:
 MgCl2: 2.5 mM
 KCl: 50 mM
 Tris-HCl: 10 mM
 pH: 8.3

FEATURES
 source 1. 668
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="8"
 /clone_lib="Human"
 STS 31. .308
 primer_bind 31. .53
 primer_bind complement(286. .308)
 BASE COUNT 230 a 104 c 113 g 221 t
 ORIGIN

Query Match 63.6%; Score 14; DB 11; Length 668;
 Best Local Similarity 63.6%; Pred. No. 1.2e+03;
 Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Oy 1 tgttgannnnnnnttgata 22

Db 543 TGTGGAAATTAACCTTGATA 564

RESULT 4
SRSSRSV/c 1015 bp RNA VRL 08-FEB-1999
LOCUS Small round structured virus genomic RNA, 3' terminal sequence
DEFINITION containing ORF2 and ORF3.
ACCESSION D38547
VERSION D38547.1 GI:560054
KEYWORDS

ORGANISM
Small round structured virus (isolate:patient; SRSV-CHIBA-407/87/J)
cDNA to genomic RNA.
small round structured virus
viruses; ssRNA positive-strand viruses, no DNA stage;
Caliciviridae; Norwalk-like viruses.

REFERENCE
1 (bases 1 to 1015)
Utogawa, E.T.
Direct Submission
Submitted (17-OCT-1994) to the DDBJ/EMBL/GenBank databases. Etsuko
T. Utogawa, National Institute of Health: 1-23-1 Toyama,
Shinjuku-ku, Tokyo 162, Japan (tel:03-5285-1111(ex.2562),
fax:03-5285-1177)
2 (bases 1 to 1015)
Utogawa, E.T., Takeda, N., Inouye, S., Kasuga, K. and Yamazaki, S.
3'-terminal sequence of a small round structured virus (SRSV) in
Japan
Arch. Virol. 135 (1-2), 185-192 (1994)
94256839

JOURNAL
MEDLINE
FEATURES
source
1. 1015
/organism="small round structured virus"
/isolate="patient; SRSV-CHIBA-407/87/J"
/db_xref="taxon:37141"
/note="SRSVs are currently recognised as a potential human
calicivirus."
<1. 298
/note="ORF2"
/codon_start=2
/protein_id="BA07549.1"
/db_xref="GI:1232112"
/translation="VPCLLPEVITHTISEQAPIQGEALLHYVDPDTRNLGEFKLY
PGGYTCVPSNSSSTGPQPLPLDGVFVAFASWVSREYQLKPVGTAGPARQLGVR"
298. 924
/note="ORF3"
/codon_start=1
/protein_id="BA07550.1"
/db_xref="GI:1232113"
/translation="MAQAIIICAIASAASAGSLGAGIQAGAEALQAOQRYQODLTLOON
SFNHDKEMLGYOMESNNLLAKNLNTRYSLLQAGSLSSDPAARVAGAPVTRLYDMG
VRVAPOSATTLRSNGNMAVPLPAOPKOKLASGYSNPYDPVORASVWQSQNSS
RSMSPYRHOALQYWTVPPTSGTSSSVSTARRGTFNDRPLFANLR"

CDS
/note="ORF2"
/codon_start=2
/protein_id="BA07549.1"
/db_xref="GI:1232112"
/translation="VPCLLPEVITHTISEQAPIQGEALLHYVDPDTRNLGEFKLY
PGGYTCVPSNSSSTGPQPLPLDGVFVAFASWVSREYQLKPVGTAGPARQLGVR"
298. 924
/note="ORF3"
/codon_start=1
/protein_id="BA07550.1"
/db_xref="GI:1232113"
/translation="MAQAIIICAIASAASAGSLGAGIQAGAEALQAOQRYQODLTLOON
SFNHDKEMLGYOMESNNLLAKNLNTRYSLLQAGSLSSDPAARVAGAPVTRLYDMG
VRVAPOSATTLRSNGNMAVPLPAOPKOKLASGYSNPYDPVORASVWQSQNSS
RSMSPYRHOALQYWTVPPTSGTSSSVSTARRGTFNDRPLFANLR"

BASE COUNT 229 a 251 c 233 g 302 t
ORIGIN

Query Match 63.6%; Score 14; DB 14; Length 1015;
Best Local Similarity 63.6%; Pred. No. 1.1e+03;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Oy 1 tgttgannnnnnnttgata 22
||||| |||||
Db 481 TGTGGACATTTCATTGATA 460

RESULT 5
BTAJ9225 1257 bp DNA MAM 05-JAN-2001
LOCUS BTAJ9225 1257 bp DNA MAM 05-JAN-2001
DEFINITION Bos taurus mariner related transposon Hamari, clone btmlec4.
ACCESSION AJ009225
VERSION AJ009225.1 GI:9187451
KEYWORDS transposon.

SOURCE
ORGANISM
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
REFERENCE
1 (bases 1 to 1257)
Demattei, M.V., Auge-Gouillou, C., Pollet, N., Hamelin, M.H.,
Meunier-Rotival, M. and Bigot, Y.
Features of the mammal marl transposons in the human, sheep, cow,
and mouse genomes and implications for their evolution
Mamm. Genome 11 (12), 1111-1116 (2000)
21015409
JOURNAL
MEDLINE
REFERENCE
1 (bases 1 to 1257)
Bigot, Y.
Direct Submission
Submitted (04-JUN-1998) Bigot Y., Institut de Recherche sur la
Biologie de l'Insecte, UPRESA CNRS 6035, Facult des Sciences, Parc
Grandmont, 37200 Tours, FRANCE
Location/Qualifiers
1. 1257
/organism="Bos taurus"
/transposon="Hamari"
/db_xref="taxon:9913"
/clone="btmlec4"

FEATURES
source
1. 1257
/organism="Bos taurus"
/transposon="Hamari"
/db_xref="taxon:9913"
/clone="btmlec4"

BASE COUNT 388 a 263 c 263 g 343 t
ORIGIN

Query Match 63.6%; Score 14; DB 4; Length 1257;
Best Local Similarity 63.6%; Pred. No. 1.1e+03;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Oy 1 tgttgannnnnnnttgata 22
||||| |||||
Db 36 TGTGGATTTCATTGATA 57

RESULT 6
HSAJ9229 1257 bp DNA PRI 06-JAN-2001
LOCUS HSAJ9229 1257 bp DNA PRI 06-JAN-2001
DEFINITION Homo sapiens mariner related transposon Hamari, clone mlehsc3.
ACCESSION AJ009229
VERSION AJ009229.1 GI:9187522
KEYWORDS transposon.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 1257)
Demattei, M.V., Auge-Gouillou, C., Pollet, N., Hamelin, M.H.,
Meunier-Rotival, M. and Bigot, Y.
Features of the mammal marl transposons in the human, sheep, cow,
and mouse genomes and implications for their evolution
Mamm. Genome 11 (12), 1111-1116 (2000)
21015409
JOURNAL
MEDLINE
REFERENCE
2 (bases 1 to 1257)
Bigot, Y.
Direct Submission
Submitted (04-JUN-1998) Bigot Y., Institut de Recherche sur la
Biologie de l'Insecte, UPRESA CNRS 6035, Facult des Sciences, Parc
Grandmont, 37200 Tours, FRANCE
Location/Qualifiers
1. 1257
/organism="Homo sapiens"
/transposon="Hamari"
/db_xref="taxon:9606"
/clone="mlehsc3"

BASE COUNT 381 a 263 c 264 g 349 t
ORIGIN

Query Match 63.6%; Score 14; DB 9; Length 1257;
Best Local Similarity 63.6%; Pred. No. 1.1e+03;

Matches 14: Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 tgttgannnnnnnnttgata 22
|||||
Db 36 TGTTCGAATTGGCCATTGTGATA 57

RESULT 7
BTAJ9224 1258 bp DNA MAM 05-JAN-2001
LOCUS Bos taurus mariner related transposon Hsmar1, clone btmlec3.
ACCESSION AJ009224 GI:9187450
VERSION AJ009224.1
KEYWORDS transposon.
SOURCE cow.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 1258)
Dematei,M.V., Auge-Gouillou,C., Pollet,N., Hamelin,M.H.,
Meunier-Rotival,M. and Bigot,Y.
Features of the mammal mar1 transposons in the human, sheep, cow,
and mouse genomes and implications for their evolution
Mamm. Genome 11 (12), 1111-1116 (2000)
21015409
Bigot,Y.
Direct Submission
Submitted (04-JUN-1998) Bigot Y., Institut de Recherche sur la
Biologie de l'Insecte, UPRESA CNRS 6035, Facult des Sciences, Parc
Grandmont, 37200 Tours, FRANCE
Location/Qualifiers
1. 1258
/organism="Bos taurus"
/transposon="Hsmar1"
/db_xref="taxon:9913"
/clone="btmlec3"

BASE COUNT 386 a 262 c 266 g 344 t

ORIGIN

Query Match 63.6%; Score 14; DB 4; Length 1258;
Best Local Similarity 63.6%; Pred. No. 1.1e+03;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 tgttgannnnnnnnttgata 22
|||||
Db 36 TGTTCGAATTGGCCATTGTGATA 57

RESULT 8
BTAJ9223 1276 bp DNA MAM 05-JAN-2001
LOCUS Bos taurus mariner related transposon Hsmar1, clone btmlec2.
ACCESSION AJ009223
VERSION AJ009223.1 GI:9187449
KEYWORDS transposon.
SOURCE cow.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 1276)
Dematei,M.V., Auge-Gouillou,C., Pollet,N., Hamelin,M.H.,
Meunier-Rotival,M. and Bigot,Y.
Features of the mammal mar1 transposons in the human, sheep, cow,
and mouse genomes and implications for their evolution
Mamm. Genome 11 (12), 1111-1116 (2000)
21015409
Bigot,Y.
Direct Submission

JOURNAL Submitted (04-JUN-1998) Bigot Y., Institut de Recherche sur la
Biologie de l'Insecte, UPRESA CNRS 6035, Facult des Sciences, Parc
Grandmont, 37200 Tours, FRANCE
Location/Qualifiers
1. 1276
/organism="Bos taurus"
/transposon="Hsmar1"
/db_xref="taxon:9913"
/clone="btmlec2"

BASE COUNT 388 a 261 c 248 g 379 t

ORIGIN

Query Match 63.6%; Score 14; DB 4; Length 1276;
Best Local Similarity 63.6%; Pred. No. 1.1e+03;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 tgttgannnnnnnnttgata 22
|||||
Db 37 TGTTCGAATTGGCCATTGTGATA 58

RESULT 9
HSU52077 1287 bp DNA PRI 26-JAN-1998
LOCUS Human mariner1 transposase gene, complete consensus sequence.
DEFINITION U52077
ACCESSION U52077
VERSION U52077.1 GI:1263080
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 1287)
Robertson,H.M. and Zampano,K.L.
Molecular evolution of an ancient mariner transposon, Hsmar1, in
the human genome
Gene 205 (1-2), 203-217 (1997)
98121293
REFERENCE 2 (bases 1 to 1287)
Robertson,H.M.
TITLE Direct Submission
JOURNAL Submitted (21-MAR-1996) Hugh M. Robertson, Entomology, University
of Illinois at Urbana-Champaign, 505 S. Goodwin, Urbana, IL 61801,
USA
Location/Qualifiers
1. 1287
/organism="Homo sapiens"
/db_xref="taxon:9606"
/transposon="Hsmar1"
/note="consensus sequence based on 20 unique sequences"

FEATURES
source
repeat_region
1. 30
/rpt_type=Inverted
179..1210
/codon_start=-1
/product="mariner transposase"
/protein_id="AAC52010.1"
/db_xref="GI:1263081"
/translation="MEMMLDKQIRAIPLFEFKMKRAETTRNNINAGPGTANERT
VQWFKFKCGDESLDEERSGRPSEVDNDLAIIEADPLTTTREVAEIYNDHSTV
VRHLKQIGKVKLDKWPYHPLSENQNRREYSSILRNNEPFLDRVTGDEKIL
YDNRRRPAQWLDRDEEAPKRPENLHOKRKVYVVMASAGLIHYSFLNGETITSEKY
AQQIDENHKKLQLOPALVNRKGPILLHDNARPHVAQPTLOKLNELGEYELPYPYSP
DLSPDTDFHFKHLDNFLQGRPHNQDAENARPOEYBSRSTDRYATGINKLISRWKCK
VDCNGSYFD"

repeat_region
1256..1287
/rpt_type=Inverted
BASE COUNT 381 a 284 c 278 g 344 t

ORIGIN

Query Match 63.6%; Score 14; DB 9; Length 1287;
Best Local Similarity 63.6%; Pred. No. 1.1e+03;

Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 ttttgannnnnnnttgata 22
|||||

Db 36 TGTGGAAATTGCGCTTGATA 57

RESULT 10
AXI48809/c 1305 bp DNA PAT 08-JUN-2001
LOCUS
DEFINITION Sequence 11 from Patent WO0136625.
ACCESSION AXI48809
VERSION AXI48809.1 GI:14347333
KEYWORDS
SOURCE Streptococcus pyogenes.
ORGANISM Streptococcus pyogenes.
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
Streptococcus.

REFERENCE
AUTHORS Wright,J.A., Young,A.H. and Dugourd,D.
TITLE Antisense oligonucleotide sequences derived from groEL and groES as
inhibitors of microorganisms
JOURNAL Patent: WO 0136625-A 11 25-MAY-2001;
Genesense Technologies Inc. (CA)
FEATURES
source 1. .1305
/organism="Streptococcus pyogenes"
/db_xref="taxon:1314"

BASE COUNT 406 a 238 c 312 g 349 t
ORIGIN

Query Match 63.6%; Score 14; DB 6; Length 1305;
Best Local Similarity 63.6%; Pred. No. 1.1e+03;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 ttttgannnnnnnttgata 22
|||||

Db 661 TGTGGAGCTCCTACTTGATA 640

RESULT 11
AXI48806/c 1306 bp DNA PAT 08-JUN-2001
LOCUS
DEFINITION Sequence 8 from Patent WO0136625.
ACCESSION AXI48806
VERSION AXI48806.1 GI:14347330
KEYWORDS
SOURCE Streptococcus pyogenes.
ORGANISM Streptococcus pyogenes.
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
Streptococcus.

REFERENCE
AUTHORS Wright,J.A., Young,A.H. and Dugourd,D.
TITLE Antisense oligonucleotide sequences derived from groEL and groES as
inhibitors of microorganisms
JOURNAL Patent: WO 0136625-A 8 25-MAY-2001;
Genesense Technologies Inc. (CA)
FEATURES
source 1. .1306
/organism="Streptococcus pyogenes"
/db_xref="taxon:1314"

BASE COUNT 407 a 235 c 308 g 356 t
ORIGIN

Query Match 63.6%; Score 14; DB 6; Length 1306;
Best Local Similarity 63.6%; Pred. No. 1.1e+03;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 ttttgannnnnnnttgata 22
|||||

Db 667 TGTGGAGCTCCTACTTGATA 646

RESULT 12
LFU60594 1481 bp DNA BCT 10-SEP-1998
LOCUS
DEFINITION Leptospiira fainei 16S ribosomal RNA gene, partial sequence.
ACCESSION U60594
VERSION U60594.1 GI:1408219
KEYWORDS
SOURCE Leptospiira fainei.
ORGANISM Leptospiira fainei.
Bacteria; Spirochaetales; Leptospiraceae; Leptospira.

REFERENCE
AUTHORS Perolat,P., Chapel,R.J., Adler,B., Baranton,G., Bulach,D.M.,
Billinghurst,M.L., Letcart,M., Merien,F. and Serrano,M.S.
TITLE Leptospiira fainei sp. nov., isolated from pigs in Australia
JOURNAL Int. J. Syst. Bacteriol. 48 Pt 3, 851-858 (1998)
MEDLINE 98404550
REFERENCE 2 (bases 1 to 1481)
AUTHORS Adler,B., Chappell,R.J., Baranton,G., Bulach,D.M.,
Billinghurst,M.L., Letcart,M., Merien,F., Serrano,M.S. and
Perolat,P.

TITLE Direct Submission
JOURNAL Submitted (12-JUN-1996) Microbiology, Monash University, Wellington
Rd., Clayton, VIC 3168, Australia
FEATURES
source 1. .1481
/organism="Leptospiira fainei"
/strain="Hurstbridge"
/db_xref="taxon:48782"

BASE COUNT 391 a 335 c 439 g 314 t 2 others
ORIGIN

Query Match 63.6%; Score 14; DB 1; Length 1481;
Best Local Similarity 63.6%; Pred. No. 1.1e+03;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 ttttgannnnnnnttgata 22
|||||

Db 155 TGTGGATCACAAGATTGATA 176

RESULT 13
DDU66913/c 2321 bp DNA INV 05-SEP-1996
LOCUS
DEFINITION Dictyostellium discoideum ORF DG1040 gene, partial cds.
ACCESSION U66913
VERSION U66913.1 GI:1519537
KEYWORDS
SOURCE Dictyostellium discoideum.
ORGANISM Dictyostellium discoideum.
Eukaryota; Mycetozoa; Dictyostellidae; Dictyostellium.

REFERENCE
AUTHORS Loomis,W.F.
TITLE Direct Submission
JOURNAL Submitted (15-AUG-1996) Dept. of Biology 0322, University of
California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA
FEATURES
source 1. .2321
/organism="Dictyostellium discoideum"
/strain="AX4"
/db_xref="taxon:44689"
/cell_line="AX4"
join(370..732,813..>2321)
/note="ORF DG1040"
<370..732
/number=1
join(370..732,813..>2321)
/note="ORF DG1040; Initially derived from a plasmid
disrupted gene"

```

/codon_start=1
/protein_id="AAB07544.1"
/db_xref="GI:1519538"
/translation="MSTRKPEIDLKRSSEVNDSTINNONNNNSINNNNDNNNNKKEE
TSEYDPPNTEFNKRLSSDKFYKATWLVGVSSNNNEQTSNTISNSLSEVATIKN
LISNLTJKRSKSNSSLSQIINQVGOAEIYFRLVDSTIDKNTNTPITTSPTTSA
TSTTPTPTPTPTPTTASSTASSSSNOLKIQLEKBEFSLTKQPHFISVLCAPFGE
VNNNEFLVFSSTLKTSPSHETMLGIALSOSIDKLVREDAKFLINTLITLSQNTKS
LPENLHOLLYKFKTOETNLOLOONNOONNOONNOONNOONNOONNOONNOONNO
OOONOOHSEVTKYILOSISPPNSLALSPVNDOSGSGGGSITSKRLQONPSTVN
SLISKVSQONVPYQIISIGYSSCSASTLKDLFPOFKLTESDIAQIIVSMAELTIN
TNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NSNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
LAVYKATGVDOQGSKPPIDFLIDRWVWNSLSQIQPLKAIQISDFEPFQNSTKKRID
"
exon
813. .>2321
/number=2
misc_feature
2124
/note="site of plasmid insertion in mutant organisms;
mutants form crinkled, branching fingers; neither spores
nor stalk cells were seen"
BASE COUNT      1055 a      338 c      210 g      718 t
ORIGIN

Query Match
Best Local Similarity 63.6%; Score 14; DB 3; Length 2321;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy      1 tgttgannnnnnnttgata 22
|||||
Db      2069 TGTGGATATATTCTTGATA 2048

RESULT 14
SPGROELGN 2683 bp DNA BCT 26-SEP-1997
LOCUS
DEFINITION
S.pyogenes DNA for groEL gene.
ACCESSION X89236
VERSION X89236.1 GI:2462691
KEYWORDS
groEL gene; heat shock protein 60 (GroEL) like protein.
SOURCE
Streptococcus pyogenes.
ORGANISM
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
Streptococcus.
REFERENCE
1 (bases 1 to 2683)
AUTHORS Podbielski,A.
TITLE Direct Submission
JOURNAL Submitted (29-JUN-1995) A. Podbielski, Institut f Med
Mikrobiologie, an der Rhein Westf. Techn. Hochschule, Aachen,
Pauwelsstr Klinikum, 52057 Aachen, FRG
2 (bases 1 to 2683)
Pohl,B., Podbielski,A. and Zarges,I.
REFERENCE Unpublished
AUTHORS Related sequences M81132, M84965.
JOURNAL Location/Qualifiers
COMMENT
FEATURES
source
1. .2683
/organism="Streptococcus pyogenes"
/strain="serotype M49"
/isolate="CS101"
/db_xref="taxon:1314"
1. .1446
/gene="groEL"
<1. .1446
/gene="groEL"
/codon_start=1
/transl_table=11
/product="heat shock protein"
/protein_id="CA61520.1"
/db_xref="GI:2462692"
/db_xref="SPPREMBL:033733"
/translation="DHFEWNGAKLVSEVASKTNDIAGDGTATATVLTQALVHGKLVN
TAGANPIGIRGIEIETATATAVALKAVAPVSGKEAIAOVAASRSSEKVEYISEAM

```

```

ERVNGDVITTEESRGMETELEVVEGQFDRCYLSQYVWTDNEKNVADLENPFILIMD
KVSNIODILPLEVEVLKTNRPLLIADVDGEALPTVLNKRGTFFNVAVKAPFG
DRKRAMIDIALITGGVITTEDLGEIADATWALGOAKITVDGSTVYVSGSGSE
AIANRILICISOLETTTSPDEKIOERAKIAGGAVAYIKVCAPEFTALKKKLIED
AINATRAAVERGIVAGGATLTTEKVALEBEDDAGRIVYRALEEPYRQALN
AYEESVVIDKLKNSPAGFNAAGVEWMDIKTGIIDPVKTRSAIONAASVASLIL
TTEAVVANKPEPAPAPAMPACMDGMMGGF"
repeat_region
1509. .1564
/rpt_type=INVERTED
BASE COUNT      808 a      490 c      593 g      792 t
ORIGIN

Query Match
Best Local Similarity 63.6%; Score 14; DB 1; Length 2683;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy      1 tgttgannnnnnnttgata 22
|||||
Db      963 TGTGGAGCTCCTACTTGATA 942

RESULT 15
AXI48805 2683 bp DNA PAT 08-JUN-2001
LOCUS
DEFINITION
Sequence 7 from Patent WO0136625.
ACCESSION AXI48805
VERSION AXI48805.1 GI:14347329
KEYWORDS
Streptococcus pyogenes.
SOURCE
Streptococcus pyogenes.
ORGANISM
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
Streptococcus.
REFERENCE
1 (bases 1 to 2683)
AUTHORS Wright,J.A., Young,A.H. and Dugourd,D.
TITLE Antisense oligonucleotide sequences derived from groEL and groES as
inhibitors of microorganisms
JOURNAL Patent: WO 0136625-A 7 25-MAY-2001;
Genesense Technologies Inc. (CA)
FEATURES
source
1. .2683
/organism="Streptococcus pyogenes"
/db_xref="taxon:1314"
BASE COUNT      808 a      490 c      593 g      792 t
ORIGIN

Query Match
Best Local Similarity 63.6%; Score 14; DB 6; Length 2683;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy      1 tgttgannnnnnnttgata 22
|||||
Db      963 TGTGGAGCTCCTACTTGATA 942

```

Search completed: December 15, 2001, 03:19:57
Job time: 7687 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 03:29:44 : Search time 401.91 Seconds
(without alignments)
46.922 Million cell updates/sec

Title: US-09-380-826a-6
Perfect score: 22
Sequence: 1 tgttgannnnnnnlttgata 22

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /SIDs2/gcgdata/geneseq/NA1980.DAT:*
2: /SIDs2/gcgdata/geneseq/NA1981.DAT:*
3: /SIDs2/gcgdata/geneseq/NA1982.DAT:*
4: /SIDs2/gcgdata/geneseq/NA1983.DAT:*
5: /SIDs2/gcgdata/geneseq/NA1984.DAT:*
6: /SIDs2/gcgdata/geneseq/NA1985.DAT:*
7: /SIDs2/gcgdata/geneseq/NA1986.DAT:*
8: /SIDs2/gcgdata/geneseq/NA1987.DAT:*
9: /SIDs2/gcgdata/geneseq/NA1988.DAT:*
10: /SIDs2/gcgdata/geneseq/NA1989.DAT:*
11: /SIDs2/gcgdata/geneseq/NA1990.DAT:*
12: /SIDs2/gcgdata/geneseq/NA1991.DAT:*
13: /SIDs2/gcgdata/geneseq/NA1992.DAT:*
14: /SIDs2/gcgdata/geneseq/NA1993.DAT:*
15: /SIDs2/gcgdata/geneseq/NA1994.DAT:*
16: /SIDs2/gcgdata/geneseq/NA1995.DAT:*
17: /SIDs2/gcgdata/geneseq/NA1996.DAT:*
18: /SIDs2/gcgdata/geneseq/NA1997.DAT:*
19: /SIDs2/gcgdata/geneseq/NA1998.DAT:*
20: /SIDs2/gcgdata/geneseq/NA1999.DAT:*
21: /SIDs2/gcgdata/geneseq/NA2000.DAT:*
22: /SIDs2/gcgdata/geneseq/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	14	63.6	22	19	AAV58901	Leptospira RNA ge
2	14	63.6	22	19	AAV58897	L. finnei nucleoti
3	14	63.6	22	19	AAV58900	Leptospira nucleot
4	14	63.6	500	22	AAI15869	Probe #5802 for ge
5	14	63.6	500	22	AAI37774	Probe #6460 used t
6	14	63.6	972	20	AAK9627	Nucleic acid seque
c 7	14	63.6	1305	22	AAH56865	Antibiotic resista
c 8	14	63.6	1306	22	AAH56862	Antibiotic resista
9	14	63.6	1477	19	AAV58896	L. finnei nucleoti
10	14	63.6	1573	22	AAH47810	Human mariner tran
c 11	14	63.6	1661	20	AAH86155	Human encoding a Str

12	14	63.6	2660	21	AAK47974	Arabidopsis thalia
c 13	14	63.6	2683	22	AAH56861	S. pyogenes groEL
c 14	14	63.6	4197	16	AAO99430	B. sphaericus SLP
c 15	13	59.1	149	21	AAC28622	Human secreted pro
c 16	13	59.1	278	22	AAI71057	C. glutamicum SRT
17	13	59.1	396	20	AAK34752	DNA encoding OMP-1
18	13	59.1	430	21	AAK09647	Human secreted pro
19	13	59.1	664	22	AAK22692	Human gastric canc
20	13	59.1	828	20	AAI26735	Human gene express
21	13	59.1	1032	22	AAK23184	DNA encoding novel
c 22	13	59.1	1079	20	AAK20676	Polynucleotide seq
c 23	13	59.1	1092	22	AAK23271	DNA encoding novel
24	13	59.1	1266	22	AAK88112	Human FLEXHT-43 nu
25	13	59.1	1482	19	AAV18099	Nucleotide sequenc
c 26	13	59.1	1669	22	AAK90031	SFO ID 1. Staphyl
c 27	13	59.1	1786	18	AAK60350	MAP kinase #2 codi
c 28	13	59.1	2084	21	AAV9696	Eucalyptus grandis
c 29	13	59.1	2133	22	AAK11387	Human phospholipas
c 30	13	59.1	2190	20	AAK07102	Staphylococcus aur
c 31	13	59.1	2340	22	AAH66171	C. glutamicum codin
c 32	13	59.1	2463	22	AAK71056	C. glutamicum SRT
c 33	13	59.1	2508	18	AAK67197	Zebrafish retinoid
c 34	13	59.1	3487	22	AAK80522	Receptor #10 parti
c 35	13	59.1	3519	21	AAK294941	Human carbohydrate
c 36	13	59.1	3519	21	AAK294948	Human carbohydrate
c 37	13	59.1	4012	18	AAV74358	Staphylococcus aur
c 38	13	59.1	4254	22	AAK61065	Human polynucleoti
c 39	13	59.1	4685	22	AAI59279	Human polynucleoti
c 40	13	59.1	4803	22	AAK11574	Human phospholipas
c 41	13	59.1	4820	22	AAK11388	Human phospholipas
c 42	13	59.1	4875	22	AAK11386	Human phospholipas
c 43	13	59.1	5253	21	AAK26868	Essential Staphylo
c 44	13	59.1	5253	22	AAK08019	Staphylococcus aur
45	13	59.1	5253	22	AAK91550	Staphylococcus aur

ALIGNMENTS

RESULT 1	
AAV58901	
ID AAV58901 standard; DNA; 22 BP.	
AC AAV58901;	
XX	
DT 20-JAN-1999 (first entry)	
XX	
DE Leptospira RNA gene nucleotide sequence.	
XX	
DE Infection; pathogenic Leptospira; protective immunity; therapy;	
KW diagnosis; ss.	
KW	
XX	
OS Leptospira sp.	
XX	
PN W09840099-A1.	
XX	
PD 17-SEP-1998.	
XX	
PF 06-MAR-1998; 98WO-AU00145.	
XX	
PR 07-MAR-1997; 97AU-0005494.	
XX	
PA (AGRI-) AGRIC VICTORIA SERVICES PTY LTD.	
PA (PICR-) PTG RES & DEV CORP.	
XX	
PI Chappel RJ;	
XX	
DR WPI; 1998-520791/44.	
XX	
PT New isolated pathogenic Leptospira bacterium - useful for, e.g	
PT developing products for conferring protective immunity, and for	
PT prophylactic or therapeutic treatment	
XX	

Arabidopsis thalia
S. pyogenes groEL
B. sphaericus SLP
Human secreted pro
C. glutamicum SRT
DNA encoding OMP-1
Human secreted pro
Human gastric canc
Human gene express
DNA encoding novel
Polynucleotide seq
DNA encoding novel
Human FLEXHT-43 nu
Nucleotide sequenc
SFO ID 1. Staphyl
MAP kinase #2 codi
Eucalyptus grandis
Human phospholipas
Staphylococcus aur
C. glutamicum codin
Zebrafish retinoid
Receptor #10 parti
Human carbohydrate
Human carbohydrate
Staphylococcus aur
Human polynucleoti
Human polynucleoti
Human phospholipas
Human phospholipas
Essential Staphylo
Staphylococcus aur
Staphylococcus aur

DT	12-OCT-2001	(first entry)
XX		
DE	Probe #5802 for gene expression analysis in human cervical cell sample	

XX Probe: human: microarray: gene expression; cervical epithelial cell;
KM cervical cancer: ss.
XX
OS Homo sapiens.
PN WO200157278-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00670.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI: 2001-488901/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human cervical epithelial cells -
XX
PS Claim 25: SEQ ID No 5802; 487pp; English.
XX
CC The present invention relates to human single exon nucleic acid probes
CC (SENPs). The present sequence is one such probe. The SENPs are derived
CC from human HeLa cells. The SENPs can be used to produce a single exon
CC microarray, which can be used for measuring human gene expression in a
CC sample derived from human cervical epithelial cells. By measuring gene
CC expression, the probes are therefore useful in grading and/or staging
CC of diseases of the cervix, notably cervical cancer.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pat_sequences.
XX
SQ Sequence 500 BP; 171 A; 80 C; 92 G; 157 T; 0 other;

Query Match 63.6%; Score 14; DB 22; Length 500;
Best Local Similarity 63.6%; Pred. No. 1.1e+02;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 tgttgannnnnnnttgata 22
||||||| |||||||
Db 65 tgttgaaatttcatttgata 86

RESULT 5
AAI37774
ID AAI37774 standard; DNA; 500 BP.
XX
AC AAI37774;
XX
DT 17-OCT-2001 (first entry)
XX
DE Probe #6460 used to measure gene expression in human placenta sample.
XX
KW Probe: microarray: human; placenta; antenatal diagnosis;
KM genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX

PF 30-JAN-2001; 2001WO-US00663.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI: 2001-488897/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human placenta -
XX
PS Claim 25: SEQ ID No 6460; 654pp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENPs).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders.
XX
SQ Sequence 500 BP; 171 A; 80 C; 92 G; 157 T; 0 other;

Query Match 63.6%; Score 14; DB 22; Length 500;
Best Local Similarity 63.6%; Pred. No. 1.1e+02;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 tgttgannnnnnnttgata 22
||||||| |||||||
Db 65 tgttgaaatttcatttgata 86

RESULT 6
AAX99627/C
ID AAX99627 standard; DNA; 972 BP.
XX
AC AAX99627;
XX
DT 05-OCT-1999 (first entry)
XX
DE Nucleic acid sequence from U. urealyticum.
XX
KW Ureaplasma urealyticum; nucleic acid detection; infection; pathogen;
KM human urogenital tract; pregnancy; neonatal disease; drug therapy;
KM suppurative arthritis; ss.
XX
OS Ureaplasma urealyticum.
XX
PN WO9939007-A1.
XX
PD 05-AUG-1999.
XX
PF 29-JAN-1999; 99WO-US01972.
XX
PR 30-JAN-1998; 98US-0073189.
XX
PA (UABR-) UAB RES FOUND.
XX
PI Cassell GH, Chen EY, Glass JI, Glass JS, Heiner CR;
PI Iefkowitz E;
XX
DR WPI: 1999-469343/39.
XX
PT Detection of Ureaplasma urealyticum using novel genes, probes and
XX primers
XX

PS Claim 1; Page 82; 110pp; English.

XX The present invention provides methods for the detection and diagnosis

CC of ureaplasma urealyticum infection. It provides novel genes

CC (AAH99501-681) that can be used as a source of primers and probes for the

CC detection and/or quantification of U. urealyticum in a biological

CC sample. The probes that can be used in the method of the invention by

CC forming target:probe complex is complementary to a region selected from

CC one of the 181 nucleic acid sequences (AAH99501-681). U. urealyticum is

CC an opportunistic pathogen of the human urogenital tract that is a

CC significant cause of adverse pregnancy outcome, neonatal disease, and

CC suppurative arthritis. As the infections are commonly asymptomatic, it is

CC important to have specific and sensitive methods for detecting their

CC presence in a patient. Also, as the pathogen has no current antibiotic

CC directed specifically against it, it would be advantageous to isolate and

CC detect gene sequences which are unique to it, and utilize these as a

CC basis for diagnosis of U. urealyticum infection as well as to develop new

CC and improved drug therapies. The present invention provides such novel

CC polynucleotide sequences (AAH99501-681).

XX

SQ Sequence 972 BP; 309 A; 127 C; 98 G; 438 T; 0 other;

Query Match 63.6%; Score 14; DB 20; Length 972;

Best Local Similarity 63.6%; Pred. No. 1.2e+02;

Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 ttttgannnnnnnttgata 22

|||||

DB 931 TGTGGACTATTAGCTTGCATA 910

RESULT 7

AAH56865/C

ID AAH56865 standard; DNA; 1305 BP.

XX

AC AAH56865;

XX

DT 06-SEP-2001 (first entry)

XX

DE Antibiotic resistant S. pyogenes groEL gene partial sequence SEQ ID.11.

XX

XX Antisense oligonucleotide; groE; groEL; groES; inhibitor; growth;

KW microorganism; Escherichia coli; Streptococcus pneumoniae; diagnosis;

KW Streptococcus pyogenes; Staphylococcus aureus; Pseudomonas aeruginosa;

KW antibacterial; antiviral; antiproliferative; antisense therapy;

KW microbial infection; ds.

XX

OS Streptococcus pyogenes.

XX

PN WO200136625-A2.

XX

PD 25-MAY-2001.

XX

PF 20-NOV-2000; 2000WO-CA01347.

XX

PR 18-NOV-1999; 99US-0166249.

XX

PA (GENE-) GENESENSE TECHNOLOGIES INC.

PI Wright JA, Young AH, Dugourd D;

XX

DR WPI; 2001-355633/37.

XX

PT Novel antisense compounds targeting nucleic acid encoding groEL or

XX groES gene of microorganism, which hybridize with and inhibit

PT expression of the genes, useful to inhibit growth of microorganism

PT having the genes -

PS Disclosure; Fig 11; 110pp; English.

XX

XX The present invention specifically claims AAH56368 to AAH56832 which are

CC antisense oligonucleotides to nucleotide sequences encoding groE. More

CC generally, antisense compounds (I) comprising antisense oligonucleotides

CC of 5-50 bases targeted to a nucleotide sequence encoding groEL (heat

CC shock protein (HSP)60) (GL) and groES (HSP10) (GS) gene from a

CC microorganism, where the antisense compound is complementary to GL or

CC GS of a microorganism and specifically hybridises with and inhibits the

CC expression of GL or GS, is claimed. (1) have antibacterial, antiviral

CC and antiproliferative activities, and can be used in antisense therapy

CC and for inhibition of expression of groES or groEL. (1) are useful for

CC inhibiting expression of GL or GS in cells or tissues in vitro. (1) are

CC also useful for inhibiting the growth of a microorganism, or inhibiting

CC the expression of GL or GS gene in a microorganism (a bacterial cell or

CC a virus) having a GL or GS gene which involves administering to the

CC microorganism or to a cell infected with the microorganism, (1). (1) are

CC also useful for treating a mammalian pathological condition mediated by

CC the microorganisms which involves identifying a eukaryotic organism

CC having a pathological condition mediated by microorganisms having a GL

CC or GS gene and administering (1) such that the growth of microorganism

CC is inhibited. The antisense compounds are utilised for diagnostics,

CC therapeutics, prophylaxis and as research reagents and kits, e.g., to

CC prevent or delay microbial infections in humans. They are also useful as

CC molecular weight markers. AAH56362 to AAH56367 and AAH56833 to AAH56854

CC represent PCR primers for groE sequences which are used in the

CC exemplification of the present invention. AAH56855 to AAH56870 represent

CC groE nucleotide sequence given in the present invention.

XX

SQ Sequence 1305 BP; 406 A; 238 C; 312 G; 349 T; 0 other;

Query Match 63.6%; Score 14; DB 22; Length 1305;

Best Local Similarity 63.6%; Pred. No. 1.3e+02;

Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 ttttgannnnnnnttgata 22

|||||

DB 661 TGTGGAGCTCTACTTTGCATA 640

RESULT 8

AAH56862/C

ID AAH56862 standard; DNA; 1306 BP.

XX

AC AAH56862;

XX

DT 06-SEP-2001 (first entry)

XX

DE Antibiotic resistant S. pyogenes groEL gene partial sequence SEQ ID.8.

XX

XX Antisense oligonucleotide; groE; groEL; groES; inhibitor; growth;

KW microorganism; Escherichia coli; Streptococcus pneumoniae; diagnosis;

KW Streptococcus pyogenes; Staphylococcus aureus; Pseudomonas aeruginosa;

KW antibacterial; antiviral; antiproliferative; antisense therapy;

KW microbial infection; ds.

XX

OS Streptococcus pyogenes.

XX

PN WO200136625-A2.

XX

PD 25-MAY-2001.

XX

PF 20-NOV-2000; 2000WO-CA01347.

XX

PR 18-NOV-1999; 99US-0166249.

XX

PA (GENE-) GENESENSE TECHNOLOGIES INC.

PI Wright JA, Young AH, Dugourd D;

XX

DR WPI; 2001-355633/37.

XX

PT Novel antisense compounds targeting nucleic acid encoding groEL or

XX groES gene of microorganism, which hybridize with and inhibit

PT expression of the genes, useful to inhibit growth of microorganism

PT having the genes -

XX Disclosure: Fig 8; 110pp; English.
PS
XX
CC The present invention specifically claims AAH56368 to AAH56832 which are
CC antisense oligonucleotides to nucleotide sequences encoding groB. More
CC generally, antisense compounds (I) comprising antisense oligonucleotides
CC of 5-50 bases targeted to a nucleotide sequence encoding groEL (heat
CC shock protein (HSP)60) (GL) and groES (HSP10) (GS) gene from a
CC microorganism, where the antisense compound is complementary to GL or
CC GS of a microorganism and specifically hybridises with and inhibits the
CC expression of GL or GS, is claimed. (I) have antibacterial, antiviral
CC and antiproliferative activities, and can be used in antisense therapy
CC and for inhibition of expression of groES or groEL. (I) are useful for
CC inhibiting expression of GL or GS in cells or tissues in vitro. (I) are
CC also useful for inhibiting the growth of a microorganism, or inhibiting
CC the expression of GL or GS gene in a microorganism (a bacterial cell or
CC a virus) having a GL or GS gene which involves administering to the
CC microorganism or to a cell infected with the microorganism, (I). (I) are
CC also useful for treating a mammalian pathological condition mediated by
CC the microorganisms which involves identifying a eukaryotic organism
CC having a pathological condition mediated by microorganisms having a GL
CC or GS gene and administering (I) such that the growth of microorganism
CC is inhibited. The antisense compounds are utilised for diagnostics,
CC therapeutics, prophylaxis and as research reagents and kits, e.g., to
CC prevent or delay microbial infections in humans. They are also useful as
CC molecular weight markers. AAH56362 to AAH56367 and AAH56833 to AAH56854
CC represent PCR primers for groB sequences which are used in the
CC exemplification of the present invention. AAH56855 to AAH56870 represent
CC groB nucleotide sequence given in the present invention.
XX
XX Sequence 1306 BP; 407 A; 235 C; 308 G; 356 T; 0 other;
SQ

Query Match 63.6%; Score 14; DB 22; Length 1306;
Best Local Similarity 63.6%; Pred. No. 1.3e+02;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 tgttgannnnnnnttgata 22
||||| |||||
Db 667 TGTGGAGCTCTCTACTTGATA 646

RESULT 9
AAV58896
ID AAV58896 standard; DNA; 1477 BP.
XX
XX AAV58896;
XX
XX 20-JAN-1999 (first entry)
DT
XX
XX L. fainei nucleotide sequence.
DE
XX
XX Infection; pathogenic Leptospira; protective immunity; therapy;
KW diagnosis; ss.
XX
XX Leptospira fainei.
OS
XX
XX WO9840099-A1.
XX
XX 17-SEP-1998.
PD
XX
XX 06-MAR-1998; 98WO-AU00145.
XX
XX 07-MAR-1997; 97AU-0005494.
PR
XX (AGRI-) AGRIC VICTORIA SERVICES PTY LTD.
PA (PIGR-) PIG RES & DEV CORP.
XX
XX Chapel RJ;
XX
XX WPI; 1998-520791/44.
XX
XX New isolated pathogenic Leptospira bacterium - useful for, e.g

PT developing products for conferring protective immunity, and for
PT prophylactic or therapeutic treatment
XX
XX
PS Claim 15; Page 69-70; 94pp; English.
XX
XX This sequence represents a Leptospira DNA sequence isolated from the
CC pathogenic Leptospira (LS) bacterium of the invention. The bacterium
CC belongs to serogroup Hurstbridge or serovar hurstbridge or the species
CC L. fainei. The LS bacteria can be used for conferring protective
CC immunity against pathogenic LS bacteria in humans or animals. The
CC bacteria can also be used for prophylactic or therapeutic treatment of LS
CC infections. The DNAs and antibodies may also be used for detection and
CC diagnosis of past or present LS infection.
XX
XX Sequence 1477 BP; 390 A; 334 C; 439 G; 314 T; 0 other;
SQ

Query Match 63.6%; Score 14; DB 19; Length 1477;
Best Local Similarity 63.6%; Pred. No. 1.3e+02;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 tgttgannnnnnnttgata 22
||||| |||||
Db 154 tgttgatcacagaatttgata 175

RESULT 10
AAH47810
ID AAH47810 standard; cDNA; 1573 BP.
XX
XX AAH47810;
XX
XX 20-SEP-2001 (first entry)
DT
XX
XX Human mariner transposase 19 encoding cDNA.
DE
XX
XX Human; mariner transposase 19; anti-HIV; immunostimulatory; cytostatic;
KW antinflammatory; diagnosis; treatment; cancer; haemopathy;
KW immunological disease; ss.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH 1026..1535
FT CDS /*tag= a
FT /*product= "mariner transposase 19"
FT /*note= "CDS claimed in claim 6"
FT
XX
XX WO200146240-A1.
XX
XX 28-JUN-2001.
PD
XX
XX 18-DEC-2000; 2000WO-CN00601.
XX
XX 22-DEC-1999; 99CN-0125674.
PR
XX (BIOW-) BROWINDOW GENE DEV INC SHANGHAI.
XX
XX Mao Y, Xie Y;
XX
XX WPI; 2001-418030/44.
XX
XX P-PSDB; AAG64431.
XX
XX Human mariner transposase 19 and encoded polynucleotide, applicable in
PT diagnosis and treatment of cancer, haemopathy, HIV infection,
PT immunological diseases and inflammation
XX
XX Claim 6; Page 30-31; 40pp; Chinese.
XX
XX The present sequence is that of the human mariner transposase 19 encoding
CC cDNA with anti-HIV, immunostimulatory, cytostatic and antinflammatory
CC activity. The polypeptide and encoded polynucleotide are applicable in
CC the diagnosis and treatment of cancer, haemopathy, HIV infection,

```
CC Immunological diseases and inflammation.
XX
SQ Sequence 1573 BP; 538 A; 297 C; 304 G; 434 T; 0 other;

Query Match      63.6%; Score 14; DB 22; Length 1573;
Best Local Similarity 63.6%; Pred. No. 1.3e+02;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 tgltggannnnnnnttgata 22
   |||||
Db 415 tgltggaattgcatttgata 436

RESULT 11
AAx86155/c
ID AAx86155 standard; DNA; 1661 BP.
XX
AC AAx86155;
XX
DT 22-SEP-1999 (first entry)
XX
DE DNA encoding a Streptococcus pyogenes heat shock protein (Hsp)60-2.
XX
KW Heat shock protein; Hsp60-2; immune response; immunological carrier;
KW cancer control; tumour; sarcoma; cancer; gene therapy; ss.
XX
OS Streptococcus pyogenes.
XX
PN W09935270-A1.
XX
PD 15-JUL-1999.
XX
PF 29-DEC-1998; 98W0-CA01203.
XX
PR 31-DEC-1997; 97U0-0001737.
XX
(STRE-) STRESSGEN BIOTECHNOLOGIES CORP.
XX
PA Mizzen L, Wisniewski J;
XX
PI WPI: 1999-430397/36.
XX
DR P-PSDB; AAY23904.
XX
DR
XX
PT New nucleic acid encoding heat shock protein-60 from Streptococcus,
PT useful in vaccines, as carriers for other immunogens, as anticancer
PT agents and for diagnosis
XX
PS Claim 3; Fig 4A-B; 176pp; English.
XX
XX
CC The present sequence encodes a heat shock protein, designated Hsp60-2.
CC The protein, its fragments, variants and fusion proteins, are
CC used to elicit or enhance an immune response against Streptococcus,
CC and to elicit a similar response to a target antigen fused to the
CC protein. Unlike other immunological carriers, Hsp60 proteins are not
CC immunosuppressive so provide an increased response to any conjugated or
CC fused antigen. Also, where used for cancer control, they lack the side
CC effects associated with endotoxins. They can also be used to detect
CC specific antibodies and in treatment or prevention of tumours
CC (e.g. sarcoma or cancers of breast, ovary, prostate, lung, pancreas or
CC liver). The Hsp60 polynucleotide is used for recombinant production
CC of the protein, as a source of primers and probes for detecting
CC streptococci in standard hybridization/amplification assays, and
CC therapeutically in gene therapy vectors.
XX
SQ Sequence 1661 BP; 512 A; 308 C; 407 G; 434 T; 0 other;

Query Match      63.6%; Score 14; DB 20; Length 1661;
Best Local Similarity 63.6%; Pred. No. 1.3e+02;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 tgltggannnnnnnttgata 22
```

```
Db 1163 TGTTGAGACTCTACTTGTGATA 1142
   |||||
RESULT 12
AAC47974
ID AAC47974 standard; DNA; 2660 BP.
XX
AC AAC47974;
XX
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 55799.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135533.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
```

PR	18-JUN-1999;	99US-0139455.	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	99US-0139456.	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	99US-0139457.	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	99US-0139458.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139459.	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	99US-0139460.	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	99US-0139461.	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	99US-0139462.	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	99US-0139463.	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	99US-0139750.	PR	27-AUG-1999;	99US-0151065.
PR	18-JUN-1999;	99US-0139763.	PR	27-AUG-1999;	99US-0151066.
PR	21-JUN-1999;	99US-0139817.	PR	27-AUG-1999;	99US-0151080.
PR	22-JUN-1999;	99US-0139899.	PR	30-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	99US-0140353.	PR	31-AUG-1999;	99US-0151438.
PR	23-JUN-1999;	99US-0140354.	PR	01-SEP-1999;	99US-0151930.
PR	24-JUN-1999;	99US-0140695.	PR	07-SEP-1999;	99US-0152363.
PR	28-JUN-1999;	99US-0140823.	PR	10-SEP-1999;	99US-0153070.
PR	29-JUN-1999;	99US-0140991.	PR	13-SEP-1999;	99US-0153758.
PR	30-JUN-1999;	99US-0141287.	PR	15-SEP-1999;	99US-0154018.
PR	01-JUL-1999;	99US-0142154.	PR	16-SEP-1999;	99US-0154039.
PR	02-JUL-1999;	99US-0142055.	PR	20-SEP-1999;	99US-0154779.
PR	06-JUL-1999;	99US-0142390.	PR	22-SEP-1999;	99US-0155139.
PR	08-JUL-1999;	99US-0142803.	PR	23-SEP-1999;	99US-0155486.
PR	09-JUL-1999;	99US-0142920.	PR	24-SEP-1999;	99US-0155659.
PR	12-JUL-1999;	99US-0142977.	PR	28-SEP-1999;	99US-0156458.
PR	13-JUL-1999;	99US-0143542.	PR	29-SEP-1999;	99US-0156596.
PR	14-JUL-1999;	99US-0143624.	PR	04-OCT-1999;	99US-0157117.
PR	15-JUL-1999;	99US-0144005.	PR	05-OCT-1999;	99US-0157753.
PR	16-JUL-1999;	99US-0144085.	PR	06-OCT-1999;	99US-0157865.
PR	16-JUL-1999;	99US-0144086.	PR	07-OCT-1999;	99US-0158029.
PR	19-JUL-1999;	99US-0144325.	PR	08-OCT-1999;	99US-0158232.
PR	19-JUL-1999;	99US-0144331.	PR	12-OCT-1999;	99US-0158369.
PR	19-JUL-1999;	99US-0144332.	PR	13-OCT-1999;	99US-0159293.
PR	19-JUL-1999;	99US-0144333.	PR	13-OCT-1999;	99US-0159294.
PR	19-JUL-1999;	99US-0144334.	PR	13-OCT-1999;	99US-0159295.
PR	19-JUL-1999;	99US-0144335.	PR	14-OCT-1999;	99US-0159329.
PR	20-JUL-1999;	99US-0144352.	PR	14-OCT-1999;	99US-0159330.
PR	20-JUL-1999;	99US-0144632.	PR	14-OCT-1999;	99US-0159331.
PR	21-JUL-1999;	99US-0144884.	PR	14-OCT-1999;	99US-0159637.
PR	21-JUL-1999;	99US-0144884.	PR	14-OCT-1999;	99US-0159638.
PR	21-JUL-1999;	99US-0145086.	PR	18-OCT-1999;	99US-0159584.
PR	21-JUL-1999;	99US-0145088.	PR	21-OCT-1999;	99US-0160741.
PR	22-JUL-1999;	99US-0145085.	PR	21-OCT-1999;	99US-0160767.
PR	22-JUL-1999;	99US-0145087.	PR	21-OCT-1999;	99US-0160768.
PR	22-JUL-1999;	99US-0145089.	PR	21-OCT-1999;	99US-0160770.
PR	22-JUL-1999;	99US-0145192.	PR	21-OCT-1999;	99US-0160814.
PR	23-JUL-1999;	99US-0145145.	PR	21-OCT-1999;	99US-0160815.
PR	23-JUL-1999;	99US-0145218.	PR	22-OCT-1999;	99US-0160980.
PR	23-JUL-1999;	99US-0145224.	PR	22-OCT-1999;	99US-0160981.
PR	26-JUL-1999;	99US-0145276.	PR	22-OCT-1999;	99US-0161404.
PR	27-JUL-1999;	99US-0145913.	PR	25-OCT-1999;	99US-0161405.
PR	27-JUL-1999;	99US-0145918.	PR	25-OCT-1999;	99US-0161406.
PR	28-JUL-1999;	99US-0145919.	PR	26-OCT-1999;	99US-0161359.
PR	28-JUL-1999;	99US-0145931.	PR	26-OCT-1999;	99US-0161360.
PR	02-AUG-1999;	99US-0146386.	PR	26-OCT-1999;	99US-0161361.
PR	02-AUG-1999;	99US-0146388.	PR	28-OCT-1999;	99US-0161920.
PR	03-AUG-1999;	99US-0146389.	PR	28-OCT-1999;	99US-0161922.
PR	04-AUG-1999;	99US-0147038.	PR	28-OCT-1999;	99US-0161993.
PR	04-AUG-1999;	99US-0147204.	PR	29-OCT-1999;	99US-0162142.
PR	04-AUG-1999;	99US-0147302.			
PR	05-AUG-1999;	99US-0147192.			
PR	05-AUG-1999;	99US-0147260.			
PR	06-AUG-1999;	99US-0147303.			
PR	06-AUG-1999;	99US-0147416.			
PR	09-AUG-1999;	99US-0147493.			
PR	09-AUG-1999;	99US-0147933.			
PR	10-AUG-1999;	99US-0148171.			
PR	11-AUG-1999;	99US-0148319.			
PR	12-AUG-1999;	99US-0148341.			
PR	13-AUG-1999;	99US-0148565.			
PR	13-AUG-1999;	99US-0148684.			
PR	16-AUG-1999;	99US-0149368.			

Query Match 63.6%; Score 14; DB 21; Length 2660;
 Best Local Similarity 63.6%; Pred. 1.4e+02;
 Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 tggtagmmmmnnmttgata 22
 |||||
 Db 955 tggtagagtcgtggttgata 976

RESULT 13
 AAH56861/C
 ID AAH56861 standard; DNA; 2663 BP.

XX AAH56861;
 AC
 XX
 DT 06-SEP-2001 (first entry)
 DE
 XX
 DE S. pyogenes groEL gene partial sequence SEQ ID NO:7.
 XX
 XX Antisense oligonucleotide; groE; groEL; groES; inhibitor; growth;
 KW microorganism; Escherichia coli; Streptococcus pneumoniae; diagnosis;
 KW Streptococcus pyogenes; Streptococcus aureus; Pseudomonas aeruginosa;
 KW antibacterial; antiviral; antiproliferative; antisense therapy;
 KW microbial infection; ds.
 XX
 XX Streptococcus pyogenes.
 OS
 XX
 PN MO200136625-A2.
 PD
 XX 25-MAY-2001.
 PF 20-NOV-2000; 2000MO-CA01347.
 XX
 PR 18-NOV-1999; 99US-0166249.
 XX
 XX (GENE-) GENESENSE TECHNOLOGIES INC.
 PA
 PI Wright JA, Young AH, Dugourd D;
 XX
 XX WPI; 2001-355633/37.
 DR
 XX
 PT Novel antisense compounds targeting nucleic acid encoding groEL or
 PT groES gene of microorganism, which hybridize with and inhibit
 PT expression of the genes, useful to inhibit growth of microorganism
 PT having the genes -
 XX
 PS Disclosure; Fig 7; 110pp; English.

XX The present invention specifically claims AAH56368 to AAH56832 which are
 CC antisense oligonucleotides to nucleotide sequences encoding groE. More
 CC generally, antisense compounds (I) comprising antisense oligonucleotides
 CC of 5-50 bases targeted to a nucleotide sequence encoding groEL (heat
 CC shock protein (HSP)60) (GL) and groES (HSP10) (GS) gene from a
 CC microorganism, where the antisense compound is complementary to GL or
 CC GS of a microorganism and specifically hybridizes with and inhibits the
 CC expression of GL or GS, is claimed. (I) have antibacterial, antiviral
 CC and antiproliferative activities, and can be used in antisense therapy
 CC and for inhibition of expression of groES or groEL. (I) are useful for
 CC inhibiting expression of GL or GS in cells or tissues in vitro. (I) are
 CC also useful for inhibiting the growth of a microorganism, or inhibiting
 CC the expression of GL or GS gene in a microorganism (a bacterial cell or
 CC a virus) having a GL or GS gene which involves administering to the
 CC microorganism or to a cell infected with the microorganism, (I). (I) are
 CC also useful for treating a mammalian pathological condition mediated by
 CC the microorganisms which involves identifying a eukaryotic organism
 CC having a pathological condition mediated by microorganisms having a GL
 CC or GS gene and administering (I) such that the growth of microorganism
 CC is inhibited. The antisense compounds are utilized for diagnostics,
 CC therapeutics, prophylaxis and as research reagents and kits, e.g., to
 CC prevent or delay microbial infections in humans. They are also useful as
 CC molecular weight markers. AAH56362 to AAH56367 and AAH56833 to AAH56854
 CC represent PCR primers for groE sequences which are used in the
 CC exemplification of the present invention. AAH56855 to AAH56870 represent
 CC groE nucleotide sequence given in the present invention.
 XX
 SQ Sequence 2683 BP; 808 A; 490 C; 593 G; 792 T; 0 other;

Query Match 63.6%; Score 14; DB 22; Length 2683;
 Best Local Similarity 63.6%; Pred. No. 1.4e+02;
 Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 tgttgannnnnnnttgata 22
 |||||||
 DB 963 TGTGGAGCTCTACTTTGATA 942

RESULT 14
 AAQ99430/C
 ID AAQ99430 standard; DNA; 4197 BP.
 XX
 XX AAQ99430;
 AC
 XX
 DT 22-DEC-1995 (first entry)
 XX
 DE B. sphaericus SLP gene.
 XX
 XX Surface layer protein; SLP; fusion protein; vaccine; antigen;
 KW surface expression; epitope; ds.
 XX
 XX Bacillus sphaericus.
 OS
 XX
 FH Key Location/Qualifiers
 FT RBS 79..85
 FT /*tag= a
 FT CDS 95..3853
 FT /*tag= b
 FT sig-peptide 95..104
 FT /*tag= c
 FT mat-peptide 185..3850
 FT /*tag= d
 XX
 PN MO9519371-A2.
 PD
 XX 20-JUL-1995.
 PF 13-JAN-1995; 95MO-EP00147.
 XX
 PR 14-JAN-1994; 94GB-0000650.
 XX
 PA (SOLV) SOLVAY SA.
 PI Deblaere RX, Desomer J, Dhaese P;
 XX
 DR WPI; 1995-263827/34.
 DR P-PADB; AAR80530.
 XX
 PT Host cell expressing surface layer protein fusion protein - used for
 PT host presentation of antigens and vaccine prodn.
 XX
 PS Disclosure; Fig.6; 95pp; English.

XX A probe based on the N-terminal sequence of B. sphaericus P-1 (LMG
 CC P-13855) surface layer protein was used to screen an HindIII-
 CC generated library to isolate the slp gene. Promoter regions
 CC of the gene are used in genetic constructs providing surface
 CC expression of heterologous proteins in P-1 hosts.
 XX
 SQ Sequence 4197 BP; 1470 A; 731 C; 763 G; 1233 T; 0 other;

Query Match 63.6%; Score 14; DB 16; Length 4197;
 Best Local Similarity 63.6%; Pred. No. 1.5e+02;
 Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 tgttgannnnnnnttgata 22
 |||||||
 DB 2836 TGTGGAGCAACTGTTTGATA 2815

RESULT 15
 AAC28622
 ID AAC28622 standard; cDNA; 149 BP.
 XX
 AC AAC28622;
 XX
 DT 06-OCT-2000 (first entry)
 XX


```

DE Human secreted protein 5' EST, SEQ ID NO: 32697.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX
OS Homo sapiens.
XX
PN EP1033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-0200610.
XX
PR 26-FEB-1999; 99US-0122487.
XX
PA (GEST ) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR WPI; 2000-500381/45.
XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
PS Claim 1: SEQ ID 32697; 71pp + CD-ROM; English.
XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.
XX
SQ Sequence 149 BP; 33 A; 14 C; 37 G; 65 T; 0 other;

Query Match          59.1%; Score 13; DB 21; Length 149;
Best Local Similarity 61.9%; Pred. No. 3.4e+02;
Matches 13; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 tgttgagannnnnnnttgat 21
   |||||
DB 13 tgttgatcttatttgat 33

```

Search completed: December 15, 2001, 03:23:46
Job time: 7886 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 03:23:47 : Search time 172.39 Seconds
(Without alignments)
28.903 Million cell updates/sec

Title: US-09-380-826a-6

Perfect score: 22

Sequence: 1 tggtagannnnnnnttgata 22

Scoring table: IDENTITY-NUC

Searched: Gap0 10.0, Gapext 1.0

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA:*

1: /cgn2_6/ptodata/2/1na/5A.COMB.seq:*

2: /cgn2_6/ptodata/2/1na/5B.COMB.seq:*

3: /cgn2_6/ptodata/2/1na/6A.COMB.seq:*

4: /cgn2_6/ptodata/2/1na/6B.COMB.seq:*

5: /cgn2_6/ptodata/2/1na/PCBUS.COMB.seq:*

6: /cgn2_6/ptodata/2/1na/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysts of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	14	63.6	3666	2	US-08-682-517-13
2	14	63.6	3666	2	US-08-682-517-14
3	14	63.6	4197	2	US-08-682-517-7
4	14	63.6	4197	2	US-08-682-517-8
5	13	59.1	3487	4	US-09-276-531-32
6	13	59.1	5253	3	US-08-714-918-19
7	13	59.1	5253	4	US-09-265-315-19
8	13	59.1	5253	4	US-09-265-315-19
9	13	59.1	5253	4	US-09-266-417-19
10	12.4	56.4	1390	4	US-08-821-994-61
11	12.4	56.4	1434	4	US-08-821-994-62
12	12.4	56.4	1760	1	US-08-413-118-118
13	12.4	56.4	1760	3	US-08-473-446-118
14	12.4	56.4	1800	1	US-08-752-238-2
15	12.4	56.4	1800	3	US-09-085-603B-2
16	12.4	56.4	1800	3	US-09-031-897-6
17	12.4	56.4	2730	1	US-08-339-129-1
18	12.4	56.4	2761	1	US-08-752-238-1
19	12.4	56.4	2761	3	US-09-085-603B-1
20	12.4	56.4	2761	3	US-09-031-897-5
21	12.4	56.4	2761	4	US-09-257-770-5
22	12.4	56.4	3438	3	US-08-613-009A-1
23	12.4	56.4	3997	4	US-08-821-994-72
24	12.4	54.5	527	4	US-08-954-395A-19
25	12	54.5	1130	1	US-07-864-004B-1
26	12	54.5	1130	1	US-08-251-937A-1
27	12	54.5	1130	1	US-08-212-133A-5

28	12	54.5	1130	1	US-08-474-503-3	Sequence 3, Appl
29	12	54.5	1130	2	US-08-670-707A-3	Sequence 3, Appl
30	12	54.5	1130	4	US-09-037-601-3	Sequence 3, Appl
31	12	54.5	1130	5	PCT-US93-03275-1	Sequence 1, Appl
32	12	54.5	1130	5	PCT-US94-13200-3	Sequence 3, Appl
33	12	54.5	1202	4	US-09-293-322C-9	Sequence 9, Appl
34	12	54.5	1623	1	US-08-121-202-3	Sequence 3, Appl
35	12	54.5	2026	3	US-08-755-587-26	Sequence 26, Appl
36	12	54.5	2625	2	US-08-468-036-2	Sequence 2, Appl
37	12	54.5	2625	2	US-08-376-843-2	Sequence 2, Appl
38	12	54.5	3284	3	US-09-136-652-1	Sequence 1, Appl
39	12	54.5	4173	4	US-08-981-729-9	Sequence 9, Appl
40	12	54.5	4173	4	US-08-981-446B-2	Sequence 2, Appl
41	12	54.5	4334	2	US-08-670-707A-38	Sequence 38, Appl
42	12	54.5	4334	4	US-09-037-601-38	Sequence 20, Appl
43	12	54.5	4931	3	US-09-058-489-20	Sequence 1, Appl
44	12	54.5	6057	3	US-08-362-525-1	Sequence 1, Appl
45	12	54.5	6402	2	US-08-670-707A-36	Sequence 36, Appl

ALIGNMENTS

RESULT 1

US-08-682-517-13/C

Sequence 13, Application US/08682517

Patent No. 5874267

GENERAL INFORMATION:

APPLICANT: TITLE OF INVENTION: Expression of surface layer proteins

NUMBER OF SEQUENCES: 25

COMPUTER READABLE FORM: MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/682,517

FILING DATE:

CLASSIFICATION:

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 3666 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-682-517-13

Query Match 63.6%; Score 14; DB 2; Length 3666;

Best Local Similarity 63.6%; Pred. No. 32;

Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 tggtagannnnnnnttgata 22

Db 2652 tggtagacacactggttgata 2631

RESULT 2

US-08-682-517-14/C

Sequence 14, Application US/08682517

Patent No. 5874267

GENERAL INFORMATION:

APPLICANT: TITLE OF INVENTION: Expression of surface layer proteins

NUMBER OF SEQUENCES: 25

COMPUTER READABLE FORM: MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:

```

? APPLICATION NUMBER: US/08/682,517
? FILING DATE:
? CLASSIFICATION:
? INFORMATION FOR SEQ ID NO: 14:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 3666 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: DNA (genomic)
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 1..3666
? US-08-682-517-14

```

Query Match	63.68;	Score 14;	DB 2;	Length 3666;
Best Local Similarity	63.68;	Pred. No. 32;		
Matches 14;	Conservative 0;	Mismatches 8;	Indels 0;	Gaps 0;

```

Qy      1  cgttgannnnnnnttggata 22
          |||||
Db 2652  TGTTCGAGCACTGGTTTGATA 2631

```

```

RESULT 3
US-08-682-517-7/c
: Sequence 7, Application US/08682517
: Patent No. 5874267
: GENERAL INFORMATION:
: APPLICANT:
: TITLE OF INVENTION: Expression of surface layer proteins
: NUMBER OF SEQUENCES: 25
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/682,517
: FILING DATE:
: CLASSIFICATION:
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 4197 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
US-08-682-517-7

```

Query Match	63.68;	Score 14;	DB 2;	Length 4197;
Best Local Similarity	63.68;	Pred. No. 33;		
Matches 14;	Conservative	0;	Mismatches 8;	Indels 0;
				Gaps 0;

```
QY      1  tgttgagannnnnnnttggata 22
          |||||  |||||
Db      2836 TGTGGAGCAACTGGTTTGATA 2815
```

RESULT 4
 US-08-682-517-8/c
 ; Sequence 8, Application US/08682517
 ; Patent No. 5874267
 ;
 ; GENERAL INFORMATION:
 ;
 ; APPLICANT:
 ; TITLE OF INVENTION: Expression of surface layer proteins
 ;
 ; NUMBER OF SEQUENCES: 25
 ;
 ; COMPUTER READABLE FORM:
 ;
 ; MEDIUM TYPE: Floppy disk
 ;
 ; COMPUTER: IBM PC compatible
 ;
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ;

```

SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/682,517
FILING DATE:
CLASSIFICATION:
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 4197 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Bacillus sphaericus
INDIVIDUAL ISOLATE: P-1
FEATURE:
NAME/KEY: CDS
LOCATION: 95..3850
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 185..3850
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 95..184
US-08-682-517-8

```

Query Match	63.6%;	Score 14;	DB 2;	Length 4197;
Best Local Similarity	63.6%;	Pred. No. 33;		
Matches 14;	Conservative 0;	Mismatches 8;	Indels 0;	Gaps 0;

```

Qy      1  tgttgannnnnnnttgata  22
          |||||
Db 2836 TGTGGAGCAACTGTTTGATA 2815

```

RESULT 5
US-09-276-531-32/c
Sequence 32. Application US/09276531
Patent No. 6183968
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti
APPLICANT: Hillman, Jennifer L.
APPLICANT: Yue, Henry
APPLICANT: Reddy, Roopa
APPLICANT: Guegler, Karl J.
APPLICANT: Baughn, Mariah R.
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF GENES ENCODING
TITLE OF INVENTION: RECEPTORS AND PROTEINS ASSOCIATED WITH CELL PROLIFERATION
NUMBER OF SEQUENCES: 134
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/276,531
FILING DATE: Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/079,677
FILING DATE: March 27, 1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Lynn E. Murty, Ph.D.

REGISTRATION NUMBER: 42,918
REFERENCE/DOCKET NUMBER: PA-0008 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 3487 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: PROSTUT01
CLONE: 714029
US-09-276-531-32

Query Match 59.1%; Score 13; DB 4; Length 3487;
Best Local Similarity 61.9%; Pred. No. 1.1e+02;
Matches 13; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 tttggannnnnnnttgat 21
|||||
Db 1880 GTTGGAGTCACTGTTGAT 1860

RESULT 6
US-08-714-918-19
Sequence 19, Application US/08714918
Patent No. 6037123
GENERAL INFORMATION:
APPLICANT: Benton, Bret
APPLICANT: Lee, Ving
APPLICANT: Malouin, Francois
APPLICANT: Martin, Patrick K.
APPLICANT: Schmid, Molly B.
APPLICANT: Sun, Dongxu
TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS ANTIBACTERIAL
TITLE OF INVENTION: TARGET GENES
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/714,918
FILING DATE: September 13, 1996
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/009,102
FILING DATE: December 22, 1995
APPLICATION NUMBER: 60/003,798
FILING DATE: September 15, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 222/005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:

LENGTH: 5253 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-714-918-19

Query Match 59.1%; Score 13; DB 3; Length 5253;
Best Local Similarity 61.9%; Pred. No. 1.2e+02;
Matches 13; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 2 gttggannnnnnnttgata 22
|||||
Db 2895 GTTGGATGATGCTTTGATA 2915

RESULT 7
US-09-265-315-19
Sequence 19, Application US/09265315
Patent No. 6187541
GENERAL INFORMATION:
APPLICANT: Benton, Bret
APPLICANT: Lee, Ving J.
APPLICANT: Malouin, Francois
APPLICANT: Martin, Patrick K.
APPLICANT: Schmid, Molly B.
APPLICANT: Sun, Dongxu
TITLE OF INVENTION: METHODS OF SCREENING FOR COMPOUNDS
TITLE OF INVENTION: TARGET GENES
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/265,315
FILING DATE: March 9, 1999
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/714,918
FILING DATE: September 13, 1996
APPLICATION NUMBER: 60/009,102
FILING DATE: December 22, 1995
APPLICATION NUMBER: 60/003,798
FILING DATE: September 15, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 240/247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 5253 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-265-315-19

Query Match 59.1%; Score 13; DB 4; Length 5253;
Best Local Similarity 61.9%; Pred. No. 1.2e+02;
Matches 13; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 2 gtgganmmmmnttgata 22
||||| ||||||
Db 2895 GTGGATGTAGTCTTTTGATA 2915

RESULT 8

US-09-265-315-19
; Sequence 19, Application US/09265315
; Patent No. 6187541
; GENERAL INFORMATION:
; APPLICANT: Benton, Bret
; APPLICANT: Lee, Ving J.
; APPLICANT: Malouin, Francois
; APPLICANT: Martin, Patrick K.
; APPLICANT: Schmid, Molly B.
; APPLICANT: Sun, Dongxu
; TITLE OF INVENTION: METHODS OF SCREENING FOR COMPOUNDS
; TITLE OF INVENTION: ACTIVE ON STAPHYLOCOCCUS AUREUS
; TITLE OF INVENTION: TARGET GENES
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: Storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/265,315
; FILING DATE: March 9, 1999
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/714,918
; FILING DATE: September 13, 1996
; APPLICATION NUMBER: 60/009,102
; FILING DATE: December 22, 1995
; APPLICATION NUMBER: 60/003,798
; FILING DATE: September 15, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 240/247
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5253 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-265-315-19

Query Match 59.1%; Score 13; DB 4; Length 5253;
Best Local Similarity 61.9%; Pred. No. 1.2e+02;
Matches 13; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 2 gtgganmmmmnttgata 22
||||| ||||||
Db 2895 GTGGATGTAGTCTTTTGATA 2915

RESULT 9
US-09-266-417-19
; Sequence 19, Application US/09266417
; Patent No. 6228588
; GENERAL INFORMATION:
; APPLICANT: Benton, Bret
; APPLICANT: Lee, Ving J.
; APPLICANT: Malouin, Francois
; APPLICANT: Martin, Patrick K.
; APPLICANT: Schmid, Molly B.
; APPLICANT: Sun, Dongxu
; TITLE OF INVENTION: METHODS OF SCREENING FOR COMPOUNDS
; TITLE OF INVENTION: ACTIVE ON STAPHYLOCOCCUS AUREUS
; TITLE OF INVENTION: TARGET GENES
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: Storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/266,417
; FILING DATE: March 9, 1999
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/714,918
; FILING DATE: September 13, 1996
; APPLICATION NUMBER: 60/009,102
; FILING DATE: December 22, 1995
; APPLICATION NUMBER: 60/003,798
; FILING DATE: September 15, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 240/248
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5253 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-266-417-19

Query Match 59.1%; Score 13; DB 4; Length 5253;
Best Local Similarity 61.9%; Pred. No. 1.2e+02;
Matches 13; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 2 gtgganmmmmnttgata 22
||||| ||||||
Db 2895 GTGGATGTAGTCTTTTGATA 2915

RESULT 10
US-08-821-994-61/c
; Sequence 61, Application US/08821994A
; Patent No. 6228643
; GENERAL INFORMATION:

```

: APPLICANT: Greenland, Andrew J
: APPLICANT: Thomas, Didier RP
: TITLE OF INVENTION: Promoters
: FILE REFERENCE: PPD 50108
: CURRENT APPLICATION NUMBER: US/08/821,994A
: EARLIER FILING DATE: 1997-03-22
: EARLIER APPLICATION NUMBER: PCT/GB97/00729
: EARLIER FILING DATE: 1997-03-18
: EARLIER APPLICATION NUMBER: GB 9606062.9
: NUMBER OF SEQ ID NOS: 89
: SOFTWARE: Patentln Ver. 2.1
: SEQ ID NO 61
: LENGTH: 1390
: TYPE: DNA
: ORGANISM: Brassica napus
US-08-821-994-61

```

```

Query Match          56.4%; Score 12.4; DB 4; Length 1390;
Best Local Similarity 59.1%; Pred. No. 2.1e+02;
Matches 13; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```

```

Qy      1 ttttgagannnnnnnttgata 22
        |||||
Db      26 TGTGCTTTCTAGTTTGATA 5

```

```

RESULT 11
US-08-821-994-62/C
: Sequence 62, Application US/08821994A
: Patent No. 6228643
: GENERAL INFORMATION:
: APPLICANT: Greenland, Andrew J
: APPLICANT: Thomas, Didier RP
: APPLICANT: Jepson, Ian
: TITLE OF INVENTION: Promoters
: FILE REFERENCE: PPD 50108
: CURRENT APPLICATION NUMBER: US/08/821,994A
: CURRENT FILING DATE: 1997-03-22
: EARLIER APPLICATION NUMBER: PCT/GB97/00729
: EARLIER FILING DATE: 1997-03-18
: EARLIER APPLICATION NUMBER: GB 9606062.9
: EARLIER FILING DATE: 1996-03-22
: NUMBER OF SEQ ID NOS: 89
: SOFTWARE: Patentln Ver. 2.1
: SEQ ID NO 62
: LENGTH: 1434
: TYPE: DNA
: ORGANISM: Brassica napus
US-08-821-994-62

```

```

Query Match          56.4%; Score 12.4; DB 4; Length 1434;
Best Local Similarity 59.1%; Pred. No. 2.1e+02;
Matches 13; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```

```

Qy      1 ttttgagannnnnnnttgata 22
        |||||
Db      26 TGTGCTTTCTAGTTTGATA 5

```

```

RESULT 12
US-08-413-118-118
: Sequence 118, Application US/08413118
: Patent No. 5688920
: GENERAL INFORMATION:
: APPLICANT: PAOLETTI, ENZO
: APPLICANT: LIMBACH, KEITH J
: TITLE OF INVENTION: NOCLOETIDE AND AMINO ACID SEQUENCES OF
: CANINE HERPESVIRUS 9B, 9C, AND 9D AND USES THEREFOR
: NUMBER OF SEQUENCES: 128

```

```

CORRESPONDENCE ADDRESS:
ADDRESS: CURTIS, MORRIS & SAFORD, P.C.
STREET: 530 FIFTH AVENUE, 25TH FLOOR
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: UNITED STATES OF AMERICA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/413,118
FILING DATE: 29-MAR-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/220,151
FILING DATE: 30-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: FROMMER, WILLIAM S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2670
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
INFORMATION FOR SEQ ID NO: 118:
SEQUENCE CHARACTERISTICS:
LENGTH: 1760 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-413-118-118

```

```

Query Match          56.4%; Score 12.4; DB 1; Length 1760;
Best Local Similarity 59.1%; Pred. No. 2.2e+02;
Matches 13; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```

```

Qy      1 ttttgagannnnnnnttgata 22
        |||||
Db      1076 TGTGATAACTATTTTATTA 1097

```

```

RESULT 13
US-08-473-446-118
: Sequence 118, Application US/08473446
: Patent No. 6017542
: GENERAL INFORMATION:
: APPLICANT: PAOLETTI, ENZO
: APPLICANT: LIMBACH, KEITH J.
: TITLE OF INVENTION: NOCLOETIDE AND AMINO ACID SEQUENCES OF
: CANINE HERPESVIRUS 9B, 9C, AND 9D AND USES THEREFOR
: NUMBER OF SEQUENCES: 128
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: CURTIS, MORRIS & SAFORD, P.C.
: STREET: 530 FIFTH AVENUE, 25TH FLOOR
: CITY: NEW YORK
: STATE: NEW YORK
: COUNTRY: UNITED STATES OF AMERICA
: ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/473,446
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/413,118

```

```

; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: FROMMER, WILLIAM S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2670
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 118:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1760 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-473-446-118

Query Match          56.4%; Score 12.4; DB 3; Length 1760;
Best Local Similarity 59.1%; Pred. No. 2.2e+02;
Matches 13; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 ttttggannnnnnnnttgata 22
Db 1076 TGTGGATTAATAATTTTGTATA 1097

RESULT 14
US-08-752-238-2/C
; Sequence 2, Application US/08752238
; Patent No. 5804418
; GENERAL INFORMATION:
; APPLICANT: Lambowitz Dr., Alan M
; APPLICANT: Mohr Dr., Georg
; APPLICANT: Saidanha Dr., Roland
; APPLICANT: Matsura Dr., Manabu
; TITLE OF INVENTION: Method for Preparing Nucleotide
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CALFEE, HALTER & GRISWOLD
; STREET: 800 Superior Avenue
; CITY: Cleveland
; STATE: Ohio
; COUNTRY: US
; ZIP: 44114
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/752,238
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Goitlick, Maty E.
; REGISTRATION NUMBER: 34,829
; REFERENCE/DOCKET NUMBER: 24671/00103
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (216) 622-8458
; TELEFAX: (216) 241-0816
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1800 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1800
; US-08-752-238-2

```

```

Query Match          56.4%; Score 12.4; DB 1; Length 1800;
Best Local Similarity 59.1%; Pred. No. 2.2e+02;
Matches 13; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 ttttggannnnnnnnttgata 22
Db 139 TATTGGAATATAAATTTTGATA 118

RESULT 15
US-09-085-603B-2/C
; Sequence 2, Application US/09085603B
; Patent No. 6001608
; GENERAL INFORMATION:
; APPLICANT: Lambowitz Dr., Alan M
; APPLICANT: Mohr Dr., Georg
; APPLICANT: Saidanha Dr., Roland
; APPLICANT: Matsura Dr., Manabu
; APPLICANT: Yang Dr., Jiam
; APPLICANT: Zimmerly Dr., Steven
; APPLICANT: Guo Dr., Huatao
; APPLICANT: Beall Dr., Clifford J.
; TITLE OF INVENTION: Methods of Making an Rnp Particle
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CALFEE, HALTER & GRISWOLD LLP
; STREET: 800 Superior Avenue
; CITY: Cleveland
; STATE: Ohio
; COUNTRY: US
; ZIP: 44114
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/085,603B
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Docherly, Pamela A.
; REGISTRATION NUMBER: 40,591
; REFERENCE/DOCKET NUMBER: 24671/04000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (216) 622-8416
; TELEFAX: (216) 241-0816
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1800 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1800
; US-09-085-603B-2

Query Match          56.4%; Score 12.4; DB 3; Length 1800;
Best Local Similarity 59.1%; Pred. No. 2.2e+02;
Matches 13; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 ttttggannnnnnnnttgata 22
Db 139 TATTGGAATATAAATTTTGATA 118

```


Search completed: December 15, 2001, 03:22:49
Job time: 7564 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 02:33:44 ; Search time 4441.54 Seconds
(without alignments)
53.226 Million cell updates/sec

Title: US-09-380-826A-6

Perfect score: 22

Sequence: 1 ttttgagannnnnnnttgata 22

Scoring table:

IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 11351937 segs, 537289281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estipl:*
6: em_estida:*
7: em_estro:*
8: em_estov:*
9: em_hic:*
10: qb_estl1:*
11: qb_estl2:*
12: qb_hic:*
13: qb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inu:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rod:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	14	63.6	184	10	BB357781 BB357781
C 2	14	63.6	210	10	BE059640 sn35a01.y
C 3	14	63.6	228	11	BE956026 UI-M-BH4-
C 4	14	63.6	235	10	BB170254 BB170254
C 5	14	63.6	235	13	AZ925147
C 6	14	63.6	250	10	A1122094 uc46f10.r
C 7	14	63.6	261	10	AW879900 QV3-OT002
C 8	14	63.6	265	10	BB287901 BB287901
C 9	14	63.6	269	11	BR800082 MR1-CT002
C 10	14	63.6	293	11	BE634128 NF082F04D
C 11	14	63.6	298	10	AA515145 ng68c01.s
C 12	14	63.6	332	10	BE202756 EST402778

C 13	14	63.6	336	13	AZ037213
C 14	14	63.6	342	13	BH083614
C 15	14	63.6	349	10	AU112918
C 16	14	63.6	359	10	AV546877
C 17	14	63.6	364	10	AM226052
C 18	14	63.6	371	13	AQ311167
C 19	14	63.6	376	13	AQ844476
C 20	14	63.6	376	13	AQ102282
C 21	14	63.6	378	10	AM240466
C 22	14	63.6	386	11	BF318641
C 23	14	63.6	389	10	AA488558
C 24	14	63.6	399	11	R89220
C 25	14	63.6	410	10	AM624588
C 26	14	63.6	410	13	AQ150667
C 27	14	63.6	416	10	AA555292
C 28	14	63.6	416	13	AQ545780
C 29	14	63.6	427	10	A1662442
C 30	14	63.6	433	11	BE464046
C 31	14	63.6	436	11	BE463084
C 32	14	63.6	447	13	AZ273408
C 33	14	63.6	459	11	BE854379
C 34	14	63.6	467	10	AL370583
C 35	14	63.6	471	10	A1591928
C 36	14	63.6	476	13	AQ914606
C 37	14	63.6	481	11	T75954
C 38	14	63.6	482	13	B29252
C 39	14	63.6	485	11	BR226581
C 40	14	63.6	489	10	BE196823
C 41	14	63.6	490	10	BE448206
C 42	14	63.6	493	10	AA637362
C 43	14	63.6	496	13	AQ589820
C 44	14	63.6	497	10	BE353412
C 45	14	63.6	498	11	BF003327

ALIGNMENTS

AZ037213	RPCT-23-3
BH083614	RPCT-24-2
AU112918	AU112918
AV546877	AV546877
AM226052	ST76G08 P
AQ311167	CITBI-E1-
AQ844476	an37g11_A
AQ102282	HS_3029_A
AM240466	uq35b10.x
BF318641	ux29e12.x
AA488558	ab37h10.r
R89220	yp99a12..r1
AM624588	EST322533
AQ150667	HS_3197_B
AA555292	pk82g03
AQ545780	RPCT-11-3
A1662442	mc125a11.x
BE464046	UI-M-CG0P
BE463084	UI-M-CG0P
AZ273408	RPCT-23-1
BE854379	ux29e12.y
AL370583	MTA338F11
A1591928	mc131b03.y
AQ914606	nbeb0049C
T75954	10732_Lambd
B29252	T29J13TF TA
BR226581	uz12f09.x
BE196823	ug71a09.y
BE448206	ut60g10.y
AA637362	vu09f03.r
AQ589820	HS_2136_B
BE353412	EST353789
BF003327	EST431825

RESULT 1
BB357781/c
LOCUS BB357781 RIKEN full-length enriched, adult male corpus striatum Mus
DEFINITION musculus cDNA clone C030024G02 3', mRNA sequence.
ACCESSION BB357781
VERSION BB357781.1 GI:9069609
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 184)
Kono, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T., Carninci,
P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N.,
Hirozane, T., Hori, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M.,
Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N.,
Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C., Kusakabe, M.,
Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y.,
Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata,
Y., Shigemoto, Y., Shinagawa, A., Shitaki, T., Sogabe, Y., Sugahara, Y.,
Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tomioka, N., Toya,
T., Tsunoda, Y., Watahiki, A., Watanabe, S., Yamamura, T., Yamana, I.,
Yano, R., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino,
M., Muramatsu, M. and Hayashizaki, Y.
RIKEN Mouse ESTs (Kono, H., et al.)
Unpublished (2000)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,

AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT Contact: Chln, H
 National Institute of Mental Health
 6001 Executive Blvd., Room 7N-7190, MSC 9643, Bethesda, MD
 20892-9643, USA
 Tel: 301 443 1706
 Fax: 301 443 9890
 Email: mestr@mail.nih.gov

The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to identify it as a clone from the olfactory bulbs tissue cDNA library preparation: M.B. Soares Lab
 cDNA distribution: Researchers may obtain BMPAP cDNA clones from RESEARCH GENETICS. It should be noted that Bento Soares is generating a small number of additional specialized non-redundant arrays of BMPAP cDNAs whose availability will be considered under appropriate and limited collaborative arrangements
 Seq primer: M13 Forward
 POLYA=Yes.

FEATURES

source 1. 228
 Location/Qualifiers
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UI-M-BH4-Dav-a-11-0-UI"
 /clone_lib="NIH-BMAP_M_S5"
 /dev_stage="27-32 days"
 /lab_host="DH10B (Life Technologies)"
 /note="Vector: pT7SP-Pac (Pharmacia) with a modified polylinker. Site_1: Not I; Site_2: Eco RI; The NIH-BMAP_M_S5 library is a subtracted library of a series, ultimately derived from a mixture of individually tagged normalized libraries from ten regions of the mouse brain (cerebellum, brain stems, olfactory bulbs, hypothalamus, cortex, amygdala, basal ganglia, pineal gland, striatum, hippocampus) after a series of subtractions to reduce the representation of cDNAs from which ESTs had already been generated. For a detailed description of the library from which this clone was derived, please visit our web site at brainst.legu.wiowa.edu.
 TAG_LIB=NIH-BMAP_M_S5
 TAG_TISSUE=olfactory-bulbs
 TAG_SEQ=CAAGG"

BASE COUNT 64 a 47 c 42 g 75 t

ORIGIN

Query Match 63.6%; Score 14; DB 11; Length 228;
Best Local Similarity 63.6%; Pred. NO.: 2.1e+03;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 tgttgagannnnnnnttgata 22
 ||||||| |||||||
Db 144 TGTGGAAATCTGCTTTCATA 123

RESULT 4
LOCUS BB170254/c 235 bp mRNA EST 29-JUN-2000
DEFINITION BB170254 RIKEN full-length enriched, adult male hypothalamus Mus
ACCESSION BB170254
VERSION BB170254.1 GI:8829337
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE

AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 235)
 Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci,
 P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N.,
 Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M.,
 Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N.,
 Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M.,
 Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y.,
 Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata,
 Y., Shigemoto,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Sugihara,Y.,
 Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tomimaga,N., Toya,
 T., Tsunoda,Y., Watabiki,A., Watanabe,S., Yamamura,T., Yamanaka,I.,
 Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino,
 M., Muramatsu,M. and Hayashizaki,Y.
 RIKEN Mouse ESTs (Konno,H., et al.)
 Unpublished (2000)
TITLE Contact: Yoshihide Hayashizaki
JOURNAL Laboratory for Genome Exploration Research Group, RIKEN Genomic
COMMENT Sciences Center(GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Saito-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@sc.riken.go.jp,
 URL:<http://genome.gsc.riken.go.jp/>
 Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki,
 N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
 The mostabillization and thermocactivation of the thermolabile enzymes by
 trehalose and its application for the synthesis of full length
 cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
 Itoh,M., Kitanai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,
 Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki,
 Y. and Hayashizaki,Y.
 Automated filtration-based high-throughput plasmid preparation
 system. Genome Res. 9 (5), 463-470 (1999)
 Carninci,P. and Hayashizaki,Y.
 High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
 19-44 (1999)
 Please visit our web site (<http://genome.rtc.riken.go.jp>) for
 further details.

FEATURES

source 1. 235
 Location/Qualifiers
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="A230017D13"
 /clone_lib="RIKEN full-length enriched, adult male
 hypothalamus"
 /sex="male"
 /tissue_type="hypothalamus"
 /dev_stage="adult"
 /lab_host="DH10B"
 /note="Site_1: SalI; Site_2: BamHI. cDNA library was
 prepared and sequenced in Mouse Genome Encyclopedia
 Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in
 RIKEN. Division of Experimental Animal Research in Riken
 contributed to prepare mouse tissues. 1st strand cDNA was
 primed with a primer [5'
 GAGAGAGAGAGATCCAGAGCTTTTGTTTTGTTTTNN 3']. cDNA was
 prepared by using trehalose thermo-activated reverse
 transcriptase and subsequently enriched for full-length by
 cap-trapper. cDNA went through one round of normalization
 to Rot = 20.0 and subtraction to Rot = 459.0. Second
 strand cDNA was prepared with the primer adapter of
 sequence [5' GAGAGAGATTCGAGTTATTAATTAATCCCCCCCC
 3']. cDNA was cleaved with XhoI and BamHI. Vector: a
 modified pBluescript KS(+) after bulk excision from Lambda
 FLC I."

BASE COUNT 84 a 47 c 31 g 73 t

ORIGIN

Query Match

63.6%; Score 14; DB 10; Length 235;

Best Local Similarity 63.6%; Pred. No. 2.1e+03;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 ttttggannnnnnnnttgata 22
|||||||
Db 66 TGTGGACTTACATCTTTGATA 45

RESULT 5
A2925147 235 bp DNA GSS 01-APR-2001
LOCUS A2925147/c
DEFINITION 4910.ez32e09.s1 Saccharomyces paradoxus N17 Saccharomyces paradoxus
genomic clone 4910.ez32e09.s1, DNA sequence.

ACCESSION A2925147.1 GI:13496046
VERSION A2925147.1
KEYWORDS GSS.
SOURCE Saccharomyces paradoxus.
ORGANISM Saccharomyces paradoxus
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomyces.

REFERENCE 1 (bases 1 to 235)
AUTHORS Clifton, P.F., Hillier, L.W., Fulton, L., Graves, T., Miner, T., Gish
W.R., Waterston, R.H. and Johnson, M.,
TITLE Surveying Saccharomyces genomes to identify functional elements by
JOURNAL comparative DNA sequence analysis
COMMENT Unpublished (2001)
Contact: Johnston M
Department of Genetics
Washington University Medical School
Box 8232, 4566 Scott Ave., St. Louis, MO 63110, USA
Tel: 314 362 2735
Fax: 314 362 7835
Email: mjgenetics.wustl.edu
Class: random plasmid subclone.

FEATURES
source
1..235
/organism="Saccharomyces paradoxus"
/strain="N17"
/db_xref="taxon:27291"
/clone="4910.ez32e09.s1"
/clone_lib="Saccharomyces paradoxus N17"
/note="Random genomic sequence"

BASE COUNT 79 a 35 c 36 g 85 t
ORIGIN

Query Match 63.6%; Score 14; DB 13; Length 235;
Best Local Similarity 63.6%; Pred. No. 2.1e+03;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 ttttggannnnnnnnttgata 22
|||||||
Db 232 TGTGGACTTACATCTTTGATA 211

RESULT 6
A1122094 250 bp mRNA EST 02-SEP-1998
LOCUS A1122094
DEFINITION uc46110.t1 Soares_mammary_gland_NMLMG Mus musculus cDNA clone
IMAGE:1401067.5 Similar to SW:Y335.MYCGE P47577 HYPOTHETICAL
GTP-BINDING PROTEIN MG335., mRNA sequence.

ACCESSION A1122094
VERSION A1122094.1 GI:3522418
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 250)
AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisels, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and

TITLE Waterston, R.
JOURNAL The WashU-HMI Mouse EST Project
COMMENT Unpublished (1996)
Contact: Maria M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: mouseest@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MG1:912783
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 225.

FEATURES
source
1..250
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:1401067"
/clone_lib="Soares_mammary_gland_NMLMG"
/sex="female (lactating)"
/tissue_type="mammary gland"
/lab_host="DH10B"
/note="Vector: p773B-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from mammary
gland tissue from a lactating female, and was then primed
with a Not I - oligo(dT) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the
modified p773 vector. Library is normalized. Library
was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 48 a 79 c 75 g 48 t
ORIGIN

Query Match 63.6%; Score 14; DB 10; Length 250;
Best Local Similarity 63.6%; Pred. No. 2.1e+03;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 ttttggannnnnnnnttgata 22
|||||||
Db 211 TGTGGAAATCCCTCTTTGATA 232

RESULT 7
AM879900 261 bp mRNA EST 23-MAY-2000
LOCUS AM879900
DEFINITION OV3-OT0029-290300-134-b02 OT0029 Homo sapiens cDNA, mRNA sequence.
ACCESSION AM879900
VERSION AM879900.1 GI:8041910
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 261)
AUTHORS Dias Neto, E., Garcia Correa, R., Verjowski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
JOURNAL MEDLINE
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922

Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?file=ct2-qv3-OT0029-290300-134-b026t3-2000-03-296t4-1)
Seq primer: puc 18 forward
High quality sequence stop: 261.
Location/Qualifiers

FEATURES
source
1..261
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1ib="OT0029"
/dev_stage="Adult"
/note="Organ: ovary; Vector: puc18; Site_1: Sma1; Site_2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent Application No. 196716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT
ORIGIN
79 a 84 c 30 g 68 t

Query Match
Best Local Similarity 63.6%; Score 14; DB 10; Length 261;
Pred. No. 2.1e+03;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
OY 1 tcttgagannnnnnncttgata 22
|||||
Db 65 tcttgagactcagagcttgcata 44

RESULT 8
BB287901/c
LOCUS BB287901 RIKEN full-length enriched, 2 cells egg Mus musculus cDNA
DEFINITION clone B020022N08 3', mRNA sequence.
ACCESSION BB287901
VERSION BB287901.1 GI:8988350
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE
AUTHORS
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 265)
Konno, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T., Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Hori, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M., Iizawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N., Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C., Kusakabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y., Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata, Y., Shigemoto, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tomiaga, N., Toya, T., Tsunoda, Y., Watanabe, S., Yamamuro, T., Yamanaka, I., Yano, R., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
RIKEN Mouse ESTs (Konno, H., et al.)
Unpublished (2000)

TITLE
JOURNAL
COMMENT
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-resgsc.riken.go.jp,
URL: http://genome-gsc.riken.go.jp/
Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S., Sasaki, N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Thermolabile enzymes by

trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh, M., Katsunari, T., Akiyama, J., Shibata, K., Iizawa, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Oza, Y., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.
Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
Please visit our web site (http://genome.etc.riken.go.jp) for further details.
Location/Qualifiers

FEATURES
source
1..265
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_1ib="B020022N08"
/clone_1ib="RIKEN full-length enriched, 2 cells egg"
/tissue_type="egg"
/dev_stage="2 cells"
/lab_host="DH10B"
/note="Site_1: SalI; Site_2: BamHI. cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGAGAGAGATCGACGCTCTTTTCTTTTCTTTTCTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGAGAGATCGACGCTCTTTTCTTTTCTTTTCTTTVN 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified Bluescript KS(+) after bulk excision from Lambda FLC I."
BASE COUNT
ORIGIN
60 a 51 c 60 g 94 t

Query Match
Best Local Similarity 63.6%; Score 14; DB 10; Length 265;
Pred. No. 2.1e+03;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
OY 1 tcttgagannnnnnncttgata 22
|||||
Db 165 tcttgagaaatgaccttgcata 144

RESULT 9
BF800082/c
LOCUS BF800082 265 bp mRNA EST 12-JAN-2001
DEFINITION MRL-C10021-161000-004-f02 C10021 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF800082
VERSION BF800082.1 GI:12129071
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 265)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M. R., Nagai, M. A., da Silva, W. Jr., Zago, M. A., Bordin, S., Costa, F. F., Goldman, G. H., Carvalho, A. F., Matsukuma, A., Bala, G. S., Simpson, D. H., Brunstein, A., de Oliveira, P. S., Bucher, P., Jongeneel, C. V., O'Hare, M. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J. and Simpson, A. J.
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=MR1&c2=MR1-C10021-
161000-004-f02ct3=2000-10-16&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 12
High quality sequence stop: 269.
Location/Qualifiers

FEATURES
source
1.269
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="C10021"
/dev_stage="Adult"
/note="Organ: colon_ins; Vector: puc18; Site:1: SmaI;
Site:2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196/716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

BASE COUNT
ORIGIN
88 a 53 c 44 g 84 t

Query Match 63.6%; Score 14; DB 11; Length 269;
Best Local Similarity 63.6%; Pred. No. 2.1e+03;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Oy 1 tggtagannnnnnnttgata 22
|||||||
Db 93 tcttgcaattctctatttgata 72

RESULT 10
BF634128 293 bp mRNA EST 19-DEC-2000
LOCUS BF634128/C
DEFINITION NP082F04D1F1042 Drought Medicago truncatula cDNA clone NF082F04DT
5' mRNA sequence.
ACCESSION BF634128
VERSION BF634128.1 GI:11898286
KEYWORDS EST.
SOURCE barrel medic.
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;
Medicago.
1 (bases 1 to 293)
Torrez-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,
Flores, H.R., Inman, J.T., Weller, J.W. and May, G.D.
Expressed Sequence Tags from the Samuel Roberts Noble Foundation
Medicago truncatula drought library
Unpublished (2000)
Contact: May GD
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 221 7391
Fax: 580 221 7380
Email: gdmay@noble.org
Insert Length: 293 Std Error: 0.00
Plate: 082 row: F column: 04
Seq primer: TCACACAGGAAACAGCTATCAGC.
location/Qualifiers
1.293
/organism="Medicago truncatula"
/db_xref="taxon:3880"

/clone="NF082F04DT"
/clone_lib="Drought"
/tissue_type="plantlets"
/dev_stage="Pooled timepoints"
/note="Vector: Lambda Zap; Contains a mixture of entire
plantlets harvested in a series of days post-watering
timepoints."

BASE COUNT
ORIGIN
109 a 76 c 25 g 83 t

Query Match 63.6%; Score 14; DB 11; Length 293;
Best Local Similarity 63.6%; Pred. No. 2.2e+03;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Oy 1 tggtagannnnnnnttgata 22
|||||||
Db 119 tcttgcaattctctatttgata 98

RESULT 11
AA515145 298 bp mRNA EST 20-AUG-1997
LOCUS AA515145
DEFINITION ng68c01.s1 NCI-CGAP_Lip2 Homo sapiens cDNA clone IMAGE:93936
similar to contains Alu repetitive element; contains element MSX1
repetitive element; mRNA sequence.
ACCESSION AA515145
VERSION AA515145.1 GI:2254745
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 298)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: David B. Krizman, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/dbp/image/image.html
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 295.
Location/Qualifiers

FEATURES
source
1.298
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:93936"
/clone_lib="NCI-CGAP_Lip2"
/tissue_type="liposarcoma"
/lab_host="DH10B"
/note="Vector: pAMP10; mRNA made from liposarcoma, cDNA
made by oligo-dT priming. Non-directionally cloned.
Size-selected on agarose gel, average insert size 600 bp.
Reference: Krizman et al. (1996) Cancer Research
56:5380-5383."

BASE COUNT
ORIGIN
76 a 63 c 91 g 68 t

Query Match 63.6%; Score 14; DB 10; Length 298;
Best Local Similarity 63.6%; Pred. No. 2.2e+03;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Oy 1 tggtagannnnnnnttgata 22
|||||||

Db 5 TGTGTGAGATGATGATTGATA 26

RESULT 12

LOCUS BE202756 332 bp mRNA EST 07-SEP-2000
DEFINITION EST402778 KVI Medicago truncatula cDNA clone pKV1-3023, mRNA
SEQUENCE
ACCESSION BE202756
VERSION BE202756.1 GI:8746021

KEYWORDS

EST.

SOURCE

Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.

REFERENCE

1 (bases 1 to 332)

AUTHORS

Vandenbosch, K., Endre, G., Hur, J., Moore, J., Beremand, P., Ellis, L.,
Town, C.D., Bowman, C.L., Craven, M.B., Hansen, T.S., Holt, I.E. and
Fraser, C.M.
ESTs from roots of Medicago truncatula 24 hours after inoculation
with Sinorhizobium meliloti
Unpublished (1999)
Contact: Vandenbosch K
Department of Biology
Texas A&M University
College Station, TX 77843-3258, USA
Tel: 409 845 7707
Fax: 409 845 2891
Email: kate@mail.bio.tamu.edu
Texas A&M University: T262522e
TIGR sequence name: MTIAE967K
More information is available at:
<http://chryslie.tamu.edu/medicago>
Seq primer: SKmd (CTA GAA GTA gta GAT CC).

JOURNAL

COMMENT

TITLE

FEATURES

source

FEATURES

source

BASE COUNT

ORIGIN

1. 332
/organism="Medicago truncatula"
/cultivar="genotype A17"
/db_xref="taxon:3880"
/clone="pKV1-3023"
/clone_lib="KVI"
/tissue_type="Seedling roots"
/dev_stage="24 hours post-inoculation with Sinorhizobium
meliloti"
/lab_host="E. coli strain XL0LR"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA
was directionally ligated into the unitap XR vector from
Stratagene and packaged using Gigapack III Gold packaging
extracts. Plasmids containing cDNA inserts were excised
from the recombinant lambda-Zap phage using Ex-assist
helper phage and propagated in XL0LR cells."

Query Match

Best Local Similarity 63.6%; Score 14; DB 10; Length 332;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Matches

1 tgttgagannnnnnnnttgata 22
||||||| |||||||

Db

157 TGTGTGATGATGCGTTTGATA 178

RESULT 13

AZ037213 336 bp DNA GSS 01-MAR-2000
LOCUS
DEFINITION RPCI-23-364L15.TV RPCI-23 Mus musculus genomic clone RPCI-23-364L15
, DNA sequence.
ACCESSION AZ037213

VERSION AZ037213.1 GI:7125453
GSS.
SOURCE house mouse.
ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 336)
Zao, S., Nierman, W., Feldblyum, T., Malek, J., Shatsman, S., Akintet
, B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Krol, M., de Jong, P.,
and Fraser, C.M.
Mouse BAC End Sequences from Library RPCI-23
Unpublished (1999)
Other_GSS: RPCI-23-364L15.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
7712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org

TITLE

JOURNAL

COMMENT

Clones are derived from the mouse BAC library RPCI-23. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.bufileo.edu). Clones may be purchased from
BACPAC Resources (<http://bacpac.med.bufileo.edu/orderingframe.htm>)
or from Resea ch Genetics (info@resgen.com). BAC end page:
http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
Plate: 364 row: L column: 15
Seq primer: T7
Class: BAC ends.

FEATURES

source

Location/Qualifiers

1. 336
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-364L15"
/clone_lib="RPCI-23"
/sex="Female"
/lab_host="DH10B"
/note="Organ: Kidney/Brain; Vector: pBAC3.6; Site_1:
EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
brain genomic DNA was isolated and partially digested
with a combination of EcoRI and EcoRI Methylase. Size
selected DNA was cloned into the pBAC3.6 vector at the
EcoRI sites. The ligation products were transformed into
DH10B electrocompetent cells (BRL Life Technologies)."

BASE COUNT

ORIGIN

97 a 37 c 45 g 157 t
1 tgttgagannnnnnnnttgata 22
||||||| |||||||

Query Match

Best Local Similarity 63.6%; Score 14; DB 13; Length 336;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Matches

1 tgttgagannnnnnnnttgata 22
||||||| |||||||

Db

268 TGTGTGATGATTCCTTTGATA 289

RESULT 14

BH083614 342 bp DNA GSS 18-JUL-2001
LOCUS BH083614/C
DEFINITION RPCI-24-293E7.TJ RPCI-24 Mus musculus genomic clone RPCI-24-293E7,
DNA sequence.
ACCESSION BH083614
VERSION BH083614.1 GI:14903211

KEYWORDS

SOURCE

ORGANISM

house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 342)
Zhao, S., Nierman, W., Malek, J., Shatsman, S., Akintet, B., Levins, M.,
Tsegaye, G., Geer, K., Krol, M., Shvartsbeyn, A., Gebregorgis, E.,

TITLE Russell, D., de Jong, P., and Fraser, C.M.
JOURNAL Mouse BAC End Sequences from Library RPCI-24
COMMENT Unpublished (1998)

Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208

Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-24. For BAC library availability, please contact Pieter de Jong (pdejong@tigr.org). Clones may be purchased from BACPAC Resources (<http://www.chori.org/bacpac/orderingframe.htm>). BAC end page: http://www.tigr.org/tldb/bac_ends/mouse/bac_end_intro.html
Plate: 293 row: E column: 7
Seq primer: SP6
Class: BAC ends.

FEATURES

source location/Qualifiers
1..342
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-24-293E7"
/clone_lib="RPCI-24"
/sex="Male"
/cell_type="Spleen/Brain"
/note="Vector: pTARBAC1; Site 1: BamHI; Site 2: BamHI; RPCI-24 Mouse BAC library produced by Pieter de Jong. The library was cloned in the pTARBAC1 cloning vector at the BamHI sites using MboI partially digested male C57BL/6J DNA."
BASE COUNT 99 a 57 c 87 g 99 t
ORIGIN

Query Match 63.6%; Score 14; DB 13; Length 342;
Best Local Similarity 63.6%; Pred. No. 2.2e+03;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1 tgttgannnnnnnttgata 22
|||||
Db 158 TGTGGACCTTTGTTGATA 137

RESULT 15

LOCUS AU112918 349 bp mRNA EST 19-OCT-2000
DEFINITION AU112918 unpublished oligo-capped cDNA library Caenorhabditis elegans cDNA clone YK7486 5', mRNA sequence.

ACCESSION AU112918
VERSION AU112918.1 GI:10926485
KEYWORDS EST.

SOURCE Caenorhabditis elegans.

ORGANISM Caenorhabditis elegans.
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditiida; Rhabditoidea
; Rhabditidae; Peloderinae; Caenorhabditis.

REFERENCE 1 (bases 1 to 349)
Kohara, Y., Shin-I, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y.
and Sugano, S.

TITLE A complementary view of the C. elegans genome
JOURNAL Unpublished (2000)
COMMENT Contact: Yuji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.

FEATURES
source 1..349
location/Qualifiers
/organism="Caenorhabditis elegans"
/strain="N2"

/db_xref="taxon:6239"
/clone="YK74866"
/clone_lib="unpublished oligo-capped cDNA library"
/sex="Hermaphrodite"
/tissue_type="whole animal"
/dev_stage="varied"
BASE COUNT 98 a 55 c 81 g 115 t
ORIGIN

Query Match 63.6%; Score 14; DB 10; Length 349;
Best Local Similarity 63.6%; Pred. No. 2.2e+03;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1 tgttgannnnnnnttgata 22
|||||
Db 279 TGTGGATATCGATTTTGTGATA 300

Search completed: December 15, 2001, 02:33:48
Job time: 4988 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 03:19:43 ; Search time 2725.73 Seconds
(without alignments)
133.153 Million cell updates/sec

Title: US-09-380-826A-2
Perfect score: 22
Sequence: 1 tttgtgacacagattgata 22

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 1472140 segs, 8248589755 residues
Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sy:*
12: gb_un:*
13: gb_vl:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sy:*
27: em_un:*
28: em_vl:*
29: em_vl:*
30: em_htgo_hum:*
31: em_htgo_inv:*
32: em_htgo_rnd:*
33: em_htg_hum:*
34: em_htg_inv:*
35: em_htg_rnd:*
36: em_htg_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	22	100.0	1481	1	LFU60594	U60594 Leptospira
2	20.4	92.7	354	1	AB007014	AB007014 Spirochaeta
3	18.8	85.5	283	1	AF003953	AF003953 Leptospira
4	18.8	85.5	288	1	LIU94974	LIU94974 Leptospira
5	18.8	85.5	288	1	LIU94975	LIU94975 Leptospira
6	18.8	85.5	288	1	LIU94976	LIU94976 Leptospira
7	18.8	85.5	288	1	LIU94977	LIU94977 Leptospira
8	18.8	85.5	288	1	LIU94978	LIU94978 Leptospira
9	18.8	85.5	288	1	LIU94979	LIU94979 Leptospira
10	18.8	85.5	353	1	AB007012	AB007012 Spirochaeta
11	18.8	85.5	1012	8	ATHRN82X	ATHRN82X
12	18.8	85.5	9423	8	ATR275979	ATR275979
13	18.8	85.5	72590	8	AC003000	AC003000 Arabidops
14	18.8	85.5	121524	2	AF165146	AF165146 Homo sapi
15	18.8	85.5	129090	2	AC083961	AC083961 Homo sapi
16	18.8	85.5	182545	2	AC046176	AC046176 Homo sapi
17	18.8	85.5	200799	8	AF137379	AF137379 Nephrosel
18	18.8	85.5	200799	8	AF137379	AF137379 Nephrosel
19	18.4	83.6	1812	3	AF220067	AF220067 Drosophila
20	18.4	83.6	27365	2	AC015209	AC015209 Drosophila
21	18.4	83.6	173613	3	AC007475	AC007475 Drosophila
22	18.4	83.6	262731	3	AE003823	AE003823 Drosophila
23	17.8	80.9	1874	14	EBBPC44NCP	EBBPC44NCP
24	17.8	80.9	31214	8	SPCC830	SPCC830
25	17.8	80.9	85992	8	AB009052	AB009052 Arabidops
26	17.8	80.9	91894	2	AC026327	AC026327 Homo sapi
27	17.8	80.9	172681	9	AC026320	AC026320 Homo sapi
28	17.8	80.9	270889	2	AC055742	AC055742 Homo sapi
29	17.4	79.1	59261	8	T12M4	T12M4
30	17.4	79.1	81662	9	AB008265	AB008265 Arabidops
31	17.4	79.1	109476	9	AL513548	AL513548 Human DNA
32	17.4	79.1	109512	2	AC068144	AC068144 Homo sapi
33	17.4	79.1	148997	2	AC021015	AC021015 Homo sapi
34	17.4	79.1	187847	9	AC023155	AC023155 Homo sapi
35	17.2	78.2	316	1	AF352068	AF352068 Uncultured
36	17.2	78.2	462	1	AF175672	AF175672 Uncultured
37	17.2	78.2	835	8	AF271231	AF271231 Albino ca
38	17.2	78.2	1083	1	AF018567	AF018567 Unidentif
39	17.2	78.2	2213	8	LEP4CCOALB	LEP4CCOALB
40	17.2	78.2	3608	3	SCYLLO31C	SCYLLO31C
41	17.2	78.2	5596	3	SPU40832	SPU40832 Strongyloce
42	17.2	78.2	34496	3	U41016	U41016 Caenorhabdi
43	17.2	78.2	38655	3	U41545	U41545 Caenorhabdi
44	17.2	78.2	49311	2	F26K10	F26K10
45	17.2	78.2	60406	9	AL590550	AL590550 Human DNA

ALIGNMENTS

RESULT	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	REFERENCE	AUTHORS
1	LFU60594	Leptospira faimel 16S ribosomal RNA gene, partial sequence.	U60594	GI:1408219			Leptospira faimel.	1 (bases 1 to 1481)	Perolat,P., Chappel,R.J., Adler,B., Baranton,G., Bulach,D.M., Billinghurst,M.L., Letocart,M., Merlen,F. and Serrano,M.S. Leptospira faimel sp. nov., isolated from pigs in Australia Int. J. Syst. Bacteriol. 48 Pt 3, 851-858 (1998)			98404550	2 (bases 1 to 1481)	Adler,B., Chappel,R.J., Baranton,G., Bulach,D.M., Billinghurst,M.L., Letocart,M., Merlen,F., Serrano,M.S. and Perolat,P.

TITLE Direct Submission
JOURNAL Submitted (12-JUN-1996) Microbiology, Monash University, Wellington Rd., Clayton, VIC 3168, Australia

FEATURES
Source
1.1481
/organism="Leptospira falnei"
/strain="Hurstbridge"
/db_xref="taxon:48782"
<1..>1481
rRNA
/product="16S ribosomal RNA"
BASE COUNT 391 a 335 c 439 g 314 t 2 others
ORIGIN

Query Match 100.0%; Score 22; DB 1; Length 1481;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tgttgatcacagaattgata 22
|||||
Db 155 TGTGGATCACAGATTGATA 176

RESULT 2
AB007014 354 bp DNA BCT 13-OCT-1997
LOCUS Spirochaeta sp. 16S rRNA gene, partial sequence.
DEFINITION AB007014
ACCESSION AB007014 GI:2516255
VERSION 16S ribosomal RNA.
KEYWORDS Spirochaeta sp. (sub_species:Freshwater obligate oligotroph, strain:FO-95) DNA.
SOURCE Spirochaeta sp.
Bacteria; Spirochaetales; Spirochaetaceae; Spirochaeta.
REFERENCE 1 (bases 1 to 354)
Shin,M.-S.
Direct Submission
Submitted (05-SEP-1997) to the DDBJ/EMBL/Genbank databases. M1-Sun Shin, Laboratory of Marine Molecular Microbiology, Faculty of Agriculture, Kyoto University, Okwacho, Kitashirakawa, Sakyo-ku, Kyoto, Kyoto 606-01, Japan (E-mail:musun@kais.kyoto-u.ac.jp, Tel:075-753-6224, Fax:075-753-6226)
2 (sites)
Shin,M., Yoshinaga,T., Uchida,A. and Ishida,Y.
Phylogenetic analysis by 16S rRNA gene sequencing of Obligate Oligotrophs Isolated from the northern basin of Lake Biwa (Mesotrophic Lake).
JOURNAL Unpublished (1997)
FEATURES
Source
Location/Qualifiers
1..354
/organism="Spirochaeta sp."
/strain="FO-95"
/sub_species="Freshwater obligate oligotroph"
/db_xref="taxon:28185"
<1..>354
rRNA
/product="16S rRNA"
BASE COUNT 91 a 79 c 111 g 73 t
ORIGIN

Query Match 92.7%; Score 20.4; DB 1; Length 354;
Best Local Similarity 95.5%; Pred. No. 13;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 tfttgatcacagaattgata 22
|||||
Db 133 TGTGGATCACAGATCTGATA 154

RESULT 3
AF003953 283 bp DNA BCT 30-MAY-1998
LOCUS Leptospira interrogans strain 48/95 16S ribosomal RNA gene, partial
DEFINITION

sequence.
ACCESSION AF003953.1 GI:3169306
VERSION AF003953.1
KEYWORDS Leptospira interrogans.
SOURCE Leptospira interrogans.
ORGANISM Bacteria; Spirochaetales; Leptospiraceae; Leptospira.
REFERENCE 1 (bases 1 to 283)
Woo,T.H.S., Patel,B.K.C., Cinco,M., Smythe,L.D., Symonds,M., Norris,M. and Dohnt,M.
Identification of Leptoneima by real-time homogeneous assay of rapid cycle PCR product
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 283)
Woo,T.H.S., Patel,B.K.C., Cinco,M., Smythe,L.D., Symonds,M., Norris,M. and Dohnt,M.
Direct Submission
Submitted (14-MAY-1997) School of Science, Griffith University, Brisbane, QLD 4111, Australia
LOCATION/Qualifiers
1..283
/organism="Leptospira interrogans"
/strain="48/95"
/db_xref="taxon:173"
<1..>283
rRNA
/product="16S rRNA"
BASE COUNT 74 a 60 c 88 g 61 t
ORIGIN

Query Match 85.5%; Score 18.8; DB 1; Length 283;
Best Local Similarity 90.9%; Pred. No. 78;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 tgttgatcacagaattgata 22
|||||
Db 136 TATTGATCACAGATTGATA 157

RESULT 4
U094974 288 bp DNA BCT 01-JAN-1998
LOCUS Leptospira inada 16S ribosomal RNA gene, partial sequence.
DEFINITION U094974
ACCESSION U094974 GI:2735446
VERSION 16S ribosomal RNA.
KEYWORDS Leptospira inada.
SOURCE Leptospira inada.
Bacteria; Spirochaetales; Leptospiraceae; Leptospira.
REFERENCE 1 (bases 1 to 288)
Woo,T.H.S., Smythe,L.D., Symonds,M., Norris,M., Dohnt,M., Brenner,D.J. and Patel,B.K.C.
Identification of Leptospira inada by continuously monitoring fluorescence during rapid cycle PCR
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 288)
Woo,T.H.S., Smythe,L.D., Symonds,M., Norris,M., Dohnt,M., Brenner,D.J. and Patel,B.K.C.
Direct Submission
Submitted (24-MAR-1997) School of Science, Griffith University, Brisbane, QLD 4111, Australia
LOCATION/Qualifiers
1..288
/organism="Leptospira inada"
/strain="79/95"
/db_xref="taxon:29506"
<1..>288
rRNA
/product="16S ribosomal RNA"
BASE COUNT 77 a 61 c 88 g 62 t
ORIGIN

Query Match 85.5%; Score 18.8; DB 1; Length 288;

Best Local Similarity	90.98;	Pred. No. 78;
Matches	20; Conservative	0; Mismatches
		2; Indels
		0; Gaps
		0

```
Oy      1 tgttgatcacaaattgata 22  
          | ||||| ||||| |||||  
Db     141 TATTGGATCACAGGATTGATA 162
```

RESULT 5

LOCUS	LI094975	288 bp	DNA	BCD	01-JAN-1998
DEFINITION	Leptospira inadai 16S ribosomal RNA gene, partial sequence.				
ACCESSION	U94975				
VERSION	U94975.1	GI:2735447			
ORIGIN					
ORIGIN					

ORGANISM *Leptospira inadai*
Bacteria: Spirochaetales; Leptospiaceae; *Leptospira*.
Characteristics: 1.3-2.0 μ

REFERENCE	TITLE
1 (pages 1 to 289)	
Woo, T.H.S., Smythe, L.D., Symonds, M., Norris, M., Dohnt, M., Brenner, D.J. and Patel, B. K.C.	Identification of <i>Leptospira</i> <i>inadai</i> by continuously monitoring fluorescence during rapid cycle PCR

JOURNAL Unpublished
REFERENCE 2 (pages 1 to 288)
AUTHORS Moo, T. H. S., Smythe, L. D., Symonds, M., Norris, M., Dohnt, M.,
Brenner, D. J. and Patel, B. K. C.

JOURNAL Submitted (24-MAR-1997) School of Science, Griffith University
Brisbane, QLD 4111, Australia
FEATURES Location/Qualifiers

```

/organism="Leptospira inadai"
/strain="68/94"
/db_xref="taxon:29506"
<1. >288
rRNA

```

BASE COUNT	77 a	61 c	88 g	62 t
ORIGIN				

Query Match	85.5%;	Score 18.8;	DB 1;	Length 288;
Best Local Similarity	90.9%;	Pred. No. 78;		
Matches 20; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0

```
QY      1  tgttgatcacaaatttgata 22
          | | | | | | | | | | | | | |
Db     141 TATTGATCACAGATTTGATA 162
```

RESULT 6

LOCUS	U94976	288 bp	DNA	BC1	01-JAN-1998
DEFINITION	Leptospira inadai 16S ribosomal RNA gene, partial sequence.				
ACCESSION	U94976				
VERSION	U94976.1	GI:2735448			

SOURCE	ORGANISM
	<i>Leptospira inadai</i> .
	<i>Leptospira inadai</i>
	Bacteria; Spirochaetales; Leptospiaceae; <i>Leptospira</i> .

AUTHORS	TITLE
Woo, T. H. S., Smythe, L. D., Symonds, M., Norris, M., Dohnt, M., Brenner, D. J. and Patel, B. K. C.	Identification of <i>Leptospira</i> in sand by continuously monitoring fluorescence during rapid cycle PCR

Journal
Reference
Authors

Unpublished
2 (bases 1
to 288)

Woo, T. H. S., Smythe, L. D., Symonds, M., Norris, M., Dohnt, M.
Brenner, D. J. and Patel, B. K. C.

TITLE Direct Submission
JOURNAL Submitted (24-MAR-1997) School of Science, Griffith University
Brisbane, QLD 4111, Australia
FEATURES Location/Qualifiers

source	1. .288
--------	---------

```

/db_xref="taxon:29506"
<1. .>288
rRNA

```

```

BASE COUNT      77 a      61 c      88 g      62 t
ORIGIN          /product="16S ribosomal RNA"

```

Query Match	85.5%;	Score 18.8;	DB 1;	Length 288;
Best Local Similarity	90.9%;	Pred. No. 78;		
Matches 20;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;

Qy	1	tgttgatcacagaatttgata	22
Db	141	TATTGGATCACAGGATTTGATA	162

RESULT 7

LOCUS	J1094977	288 bp	DNA	BCT	01-JAN-1998
DEFINITION	Leptospira inadai 16S ribosomal RNA gene, partial sequence.				
ACCESSION	U94977				
VERSION	U94977.1	GI:2735449			

SOURCE ORGANISM	Leptospira inadai
Leptospira inadai	
Bacteria; Spirochaetales; Leptospiaceae; Leptospira.	

AUTHORS	TITLE
Woo, T. H. S., Smythe, L. D., Symonds, M., Norris, M., Dohnt, M., Brenner, D. J. and Patel, B. K. C.	Identification of <i>Leptospira</i> <i>in</i> <i>indai</i> by continuously monitoring fluorescence during rapid cycle PCR

JOURNAL REFERENCE AUTHORS	TITLE
2 (pages 1 to 288) Woo, T. H. S., Smythe, L. D., Symonds, M., Norris, M., Dohnt, M., Brenner, D. J. and Patel, B. K. C.	Direct Submission

JOURNAL	Submitted (24-MAR-1997) School of Science, Griffith University Brisbane, QLD 4111, Australia
FEATURES	Location/Qualifiers
SOURCE	1..288

```

/organism="Leptospira inadai"
/strain="268/95"
/db_xref="taxon:29506"
<1..>288
rRNA

```

```

/product="16S ribosomal RNA"
BASE COUNT      77 a      61 c      88 g      62 t
ORIGIN

```

Query Match	85.5%	Score 18.8;	DB 1;	Length 288;
Best Local Similarity	90.9%;	Pred. No. 78;		
Matches 20;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0

OY 1 tgttgatcacaagatttgata 22
 | ||||| | | | | |
Db 141 TATTGATCACAGGATTTGATA 167

RESULT

LOCUS	LIU94978	288 bp	DNA	BCI	01-JAN-1998
DEFINITION	Leptospira inadai 16S ribosomal RNA gene, partial sequence.				
ACCESSION	U94978				
VERSION	U94978.1	GI:2735450			

KEYWORDS
SOURCE
ORGANISM

Leptospira inadai.
Leptospira inadai
Bacteria: Spirochaetales; Leptospiraceae; Leptospira

REFERENCE AUTHORS	1 (bases 1 to 288)
Woo, T. H. S., Smythe, L. D., Symonds, M., Norris, M., Dohnt, M.	

TITLE Brenner, D.J. and Patel, B.K.C.
 JOURNAL Identification of *Leptospira inadai* by continuously monitoring
 REFERENCE fluorescence during rapid cycle PCR
 JOURNAL Unpublished
 AUTHORS 2 (bases 1 to 288)
 TITLE Woo, T.H.S., Smythe, L.D., Symonds, M., Norris, M., Dohnt, M.,
 Brenner, D.J. and Patel, B.K.C.
 JOURNAL Direct Submission
 SUBMITTED (24-MAR-1997) School of Science, Griffith University,
 Brisbane, QLD 4111, Australia
 FEATURES Location/Qualifiers
 source 1..288
 /organism="Leptospira inadai"
 /strain="218/95"
 /db_xref="taxon:29506"
 <1..>288
 /product="16S ribosomal RNA"
 BASE COUNT 77 a 61 c 88 g 62 t
 ORIGIN

Query Match 85.5%; Score 18.8; DB 1; Length 288;
 Best Local Similarity 90.9%; Pred. No. 78;
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 tgttgatcacagaattgata 22
 1 ||||| ||||| ||||| |||||
 Db 141 TATTGGATCACAGATTGATA 162

RESULT 9
 LIU94979 288 bp DNA BCT 01-JAN-1998
 LOCUS *Leptospira inadai* 16S ribosomal RNA gene, partial sequence.
 DEFINITION U94979
 ACCESSION U94979.1 GI:2735451
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Bacteria; Spirochaetales; Leptospiraceae; Leptospira.
 REFERENCE 1 (bases 1 to 288)
 AUTHORS Woo, T.H.S., Smythe, L.D., Symonds, M., Norris, M., Dohnt, M.,
 Brenner, D.J. and Patel, B.K.C.
 TITLE Identification of *Leptospira inadai* by continuously monitoring
 JOURNAL fluorescence during rapid cycle PCR
 SUBMITTED (24-MAR-1997) School of Science, Griffith University,
 Brisbane, QLD 4111, Australia
 FEATURES Location/Qualifiers
 source 1..288
 /organism="Leptospira inadai"
 /strain="1078 VRI"
 /db_xref="taxon:29506"
 <1..>288
 /product="16S ribosomal RNA"
 BASE COUNT 77 a 61 c 88 g 62 t
 ORIGIN

Query Match 85.5%; Score 18.8; DB 1; Length 288;
 Best Local Similarity 90.9%; Pred. No. 78;
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 tgttgatcacagaattgata 22
 1 ||||| ||||| ||||| |||||
 Db 141 TATTGGATCACAGATTGATA 162

RESULT 10

AB007012
 LOCUS AB007012 353 bp DNA BCT 13-OCT-1997
 DEFINITION *Spirochaeta* sp. 16S rRNA gene, partial sequence.
 JOURNAL AB007012
 ACCESSION AB007012.1 GI:2516253
 VERSION
 KEYWORDS 16S ribosomal RNA.
 SOURCE *Spirochaeta* sp. (sub-species: Freshwater obligate oligotroph,
 strain: SO-104) DNA.
 ORGANISM
 Spirochaeta sp.
 Bacteria; Spirochaetales; Spirochaetaceae; Spirochaeta.
 REFERENCE 1 (bases 1 to 353)
 AUTHORS Shin, M.-S.
 TITLE Direct Submission
 JOURNAL Submitted (05-SEP-1997) to the DDBJ/EMBL/GenBank databases. M1-Sun
 Shin, Laboratory of Marine Molecular Microbiology, Faculty of
 Agriculture, Kyoto University, Oiwakecho, Kitashirakawa, Sakyo-ku,
 Kyoto, Kyoto 606-01, Japan (E-mail: misunekais.kyoto-u.ac.jp,
 Tel: 075-753-6224, Fax: 075-753-6226)
 2 (sites)
 Shin, M., Yoshinaga, I., Uchida, A. and Ishida, Y.
 Phylogenetic analysis by 16S rRNA gene sequencing of obligate
 oligotrophs isolated from the northern basin of Lake Biwa
 (Mesotrophic Lake)
 JOURNAL Unpublished (1997)
 FEATURES Location/Qualifiers
 source 1..353
 /organism="Spirochaeta sp."
 /strain="SO-104"
 /sub-species="Freshwater obligate oligotroph"
 /db_xref="taxon:28185"
 <1..>353
 /product="16S rRNA"
 BASE COUNT 90 a 80 c 108 g 75 t
 ORIGIN

Query Match 85.5%; Score 18.8; DB 1; Length 353;
 Best Local Similarity 90.9%; Pred. No. 77;
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 tgttgatcacagaattgata 22
 1 ||||| ||||| ||||| |||||
 Db 132 TGTGGATCACACAGATCTGATA 153

RESULT 11
 ATHRS2X
 LOCUS ATHRS2X 1012 bp mRNA PIN 30-OCT-1994
 DEFINITION *Arabidopsis thaliana* ribonuclease (RNS2) mRNA, complete cds.
 JOURNAL M98336
 ACCESSION M98336.1 GI:289209
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Arabidopsis thaliana CDNA to mRNA.
 REFERENCE 1 (bases 1 to 1012)
 AUTHORS Taylor, C.B., Bariola, P.A., delCardayre, S.B., Raines, R.T. and
 Green, P.J.
 TITLE RNS2: a senescence-associated RNase of *Arabidopsis* that diverged
 JOURNAL from the S-RNases before speciation
 MEDLINE Proc. Natl. Acad. Sci. U.S.A. 90 (11), 5118-5122 (1993)
 FEATURES Location/Qualifiers
 source 1..1012
 /organism="Arabidopsis thaliana"
 /db_xref="taxon:3702"
 /gene="RNS2"
 16..795
 /gene="RNS2"
 15..795
 /codon_start=1
 CDS

		/product="ribonuclease"	
		/protein_id="AA54A06.1"	
		/db_xref="GI:289210"	
		/translation="MASRLCILLIVACIAGFAGDVIELNRSOREEDYFAISLQWPGT	
		YCRRTRHICCSKNMACCGSDADPTOTTHGLMPDYNDGSPSCCYRSDPKKEISTLMGG	
		LEKVVPLSCGSPSCNGCGSGFWHGEMERKIGTCSSVFHEYNVFLTLINLYLKHNHV	
		TDVYOGAVYASNSSEKYPGLGIYAIONAPAIPEEVCKRAIDIRLCIFKDKPRDD	
		CVSQODLSRKSRCFKYVSLPEYFPLODEAWMLKMPFERAL"	
BASE COUNT	287 a	197 c	217 g 311 t
ORIGIN			
Query Match	85.5%	Score 18.8;	DB 8; Length 1012;
Best Local Similarity	90.9%	Ared. No. 76;	
Matches 20; Conservative	0;	Mismatches 2;	Indels 0; Gaps 0;
Oy	1	tgttgatcacaaagattgata 22	
Db	672	TGTTGATTCAACAAGATTGACA 693	
RESULT 12			
LOCUS ATHT275979	9423 bp	DNA	PLN 29-MAR-2001
DEFINITION Arabidopsis thaliana gene for GDP-mannose pyrophosphorylase, gene for vacuolar ribonuclease and gene for proline biosynthesis coding enzyme.			
VERSION AJ275979			
KEYWORDS A1275979.1 GI:13509286			
SOURCE GDP-mannose pyrophosphorylase; proline biosynthesis codifying enzyme; vacuolar ribonuclease.			
ORGANISM thale cress.			
Arabidopsis thaliana			
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots;			
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.			
1 (bases 1 to 9423)			
REFERENCE 1			
AUTHORS Avila,C.			
TITLE Genes responding to phosphate starvation placed together in Arabidopsis genome			
JOURNAL Unpublished			
REFERENCE 2 (bases 1 to 9423)			
AUTHORS Avila,C.			
TITLE Direct Submission			
JOURNAL Submitted (01-MAR-2000) Avila C., Biologia Molecular y Bioquímica,			
FACULTAD DE CIENCIAS, E-29071, MALAGA, SPAIN			
FEATURES 1..9423			
source /organism="Arabidopsis thaliana"			
exon /db_xref="taxon:3702"			
1405..1515			
number=1			
join(1405..1515,1597..1659,1748..1856,1883..2013,2099..2270)			
codon_start=1			
product="GDP-mannose pyrophosphorylase"			
protein_id="CAC35355.1"			
/db_xref="GI:13509287"			
translation="MKALLILVGFGSTRLRPLTSLPPRYLVDFANKPMILHOIEALKAV			
GVEVDVAINQPEVMENLFKDFETKLEIKTCQEPEPLCTAGPLALARKLIDGSGG			
EPFVLINSADVSEPLKEILIEFHSHGEASIMTKVDEPSKYGVVAMEESTGVEKERK			
VERKKLVGNKINAGIYLINSLVDKIELRTSIKEKTFPIIAAGOHGIVLPGFMWM			
DIOGRPYIGCLRIYLDSIRKSPAKLTSGPHIYCNVLDEFTATIGSCIILGPVAIG			
PQGVESGSVRLSRCTRMRGVIRKKHACISSITICMHSYVGOMARIENNNTIIIGDEVHYS			
DELISNGCVLPHEIKSNILKPIVM			
join(1405..1515,1597..1659,1748..1856,1883..2013,2099..2270)			
number=1			
1516..1596			
number=1			
1597..1659			
number=2			
1660..1747			
number=2			

exon	1748..1856
exon	1857..1862
intron	1863..2013
exon	2014..2098
intron	2099..2270
exon	2271..2388
exon	2389..2473
exon	2474..2517
exon	2518..2551
exon	2552..2585
exon	2586..2619
exon	2620..2653
exon	2654..2687
exon	2688..2721
exon	2722..2755
exon	2756..2789
exon	2790..2823
exon	2824..2857
exon	2858..2891
exon	2892..2925
exon	2926..2959
exon	2960..2993
exon	2994..3027
exon	3028..3061
exon	3062..3095
exon	3096..3129
exon	3130..3163
exon	3164..3197
exon	3198..3231
exon	3232..3265
exon	3266..3299
exon	3300..3333
exon	3334..3367
exon	3368..3401
exon	3402..3435
exon	3436..3469
exon	3470..3503
exon	3504..3537
exon	3538..3571
exon	3572..3605
exon	3606..3639
exon	3640..3673
exon	3674..3707
exon	3708..3741
exon	3742..3775
exon	3776..3809
exon	3810..3843
exon	3844..3877
exon	3878..3911
exon	3912..3945
exon	3946..3979
exon	3980..4013
exon	4014..4047
exon	4048..4081
exon	4082..4115
exon	4116..4149
exon	4150..4183
exon	4184..4217
exon	4218..4251
exon	4252..4285
exon	4286..4319
exon	4320..4353
exon	4354..4387
exon	4388..4421
exon	4422..4455
exon	4456..4489
exon	4490..4523
exon	4524..4557
exon	4558..4591
exon	4592..4625
exon	4626..4659
exon	4660..4693
exon	4694..4727
exon	4728..4761
exon	4762..4795
exon	4796..4829
exon	4830..4863
exon	4864..4897
exon	4898..4931
exon	4932..4965
exon	4966..4999
exon	5000..5033
exon	5034..5067
exon	5068..5101
exon	5102..5135
exon	5136..5169
exon	5170..5203
exon	5204..5237
exon	5238..5271
exon	5272..5305
exon	5306..5339
exon	5340..5373
exon	5374..5407
exon	5408..5441
exon	5442..5475
exon	5476..5509
exon	5510..5543
exon	5544..5577
exon	5578..5611
exon	5612..5645
exon	5646..5679
exon	5680..5713
exon	5714..5747
exon	5748..5781
exon	5782..5815
exon	5816..5849
exon	5850..5883
exon	5884..5917
exon	5918..5951
exon	5952..5985
exon	5986..6019
exon	6020..6053
exon	6054..6087
exon	6088..6121
exon	6122..6155
exon	6156..6189
exon	6190..6223
exon	6224..6257
exon	6258..6291
exon	6292..6325
exon	6326..6359
exon	6360..6393
exon	6394..6427
exon	6428..6461
exon	6462..6495
exon	6496..6529
exon	6530..6563
exon	6564..6597
exon	6598..6631
exon	6632..6665
exon	6666..6699
exon	6700..6733
exon	6734..6767
exon	6768..6801
exon	6802..6835
exon	6836..6869
exon	6870..6903
exon	6904..6937
exon	6938..6971
exon	6972..7005
exon	7006..7039
exon	7040..7073
exon	7074..7107
exon	7108..7141
exon	7142..7175
exon	7176..7209
exon	7210..7243
exon	7244..7277
exon	7278..7311
exon	7312..7345
exon	7346..7379

Genes were identified by a combination of three methods: Gene prediction programs including GRAIL (<http://arthur.epm.ornl.gov/pub/xgrail/>), GeneFINDER (Phil Green, University of Washington), GenScan (Chris Burge, <http://genome.stanford.edu/GENSCAN.html>), and NetPlantGene (<http://www.cbs.dtu.dk/services/NetGene2/>), searches of the complete sequence against a peptide database and plant EST databases at TIGR, and manual curation based on those analyses. Annotated genes are named to indicate the level of evidence for their annotation. Genes with similarity to other proteins are named after the database hits. Genes without significant peptide similarity but with EST similarity are named as 'unknown' proteins. Genes without protein or EST similarity, that are predicted by two or more gene prediction programs over most of their length are annotated as 'hypothetical' proteins. Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, <http://genome.wustl.edu/eddy/tRNAscan-SE/>). Simple repeats were identified by RepeatMasker (Arian Smit, <http://ftp.genome.washington.edu/RM/RepeatMasker.html>). Genes are numbered from the top to bottom of the chromosome.

We thank the CSHL/MashU/ABI consortium for sequencing BAC clones REP23, F5J6, T17A5, and T13L16, the ESSA group for sequencing clone F13D4, and Scott Jackson, Jiming Jiang, Klaus Meyer, Eric Richards and Satoshi Tabata for helpful assistance. In addition, we would like to thank the TIGR Bioinformatics Department, especially Lixin Zhou, Hanif Khalak, Michael E. Heaney, Lily Fu, Feng Liang, Jeremy Peterson, Michael Holmes, and Delwood Richardson for software and database support.

This work was supported by the National Science Foundation, Department of Energy and the US Department of Agriculture.

Address all correspondence to: atetigr.org.

FEATURES	Location/Qualifiers
source	1..72590 /organism="Arabidopsis thaliana" /cultivar="Columbia" /db_xref="taxon:3702" /chromosome="11"
misc_feature	1..>72590 /note="Sequence from clone T517"
repeat_region	complement(130..198) /rpt_family="(CAA)n"
repeat_region	complement(1000..1024) /rpt_family="POLY_A"
repeat_region	complement(1266..1291) /rpt_family="(CAAA)n"
repeat_region	complement(2759..2862) /rpt_family="POLY_A"
mRNA	complement(<4584..>5909) /gene="At2g39710"
gene	complement(<4584..>5909) /gene="At2g39710"
	/note="T517.1: similar to Gp122450121gnl PDI e3270001297341"
CDS	complement(4584..5909) /gene="At2g39710"
	/note="unknown protein"
	/codon_start=1 /protein_id="AAB87120.1"
	/db_xref="GI:2642153"
	/translation="MDSSSSSLSLSKRNFLRISVLLIFPLTECKSTSTNOTLLSLK TOKLPSSSKLSFRNVTLTAVGDDPONTISMLDTSELSMLCKKSPNLSVYF NPVSSSYSPGSPICRTPTDLPASCDPTLCHVAISVADATSTEGNLAHET FVGSYTRPGTLRGCMDSGLSSNSEDAKSTGLMNGRSLSPFNOIGFSEKSTIG SDSGFTLLGDASTSLGPIOTYPLVQSTPLPFDFVATYVQLEGTRVSKILSLK SLVYVDHGTAGQTVNDGQTQFTFLMGVYATALKNEFTTQKSVLRLLVDDPVPQGM DICVKGSTTRPNFSGLPMVSLMRGAEVMSGCKLYRVNGASEKKEVEYCTPFGN SDLGIEAFVIGHHQNVMEFDLAKSRVGAFAVNCVDLASQRLGRD"
mRNA	complement(<10039..11304) /gene="At2g39720"

gene	complement(<10039..11304) /gene="At2g39720"
CDS	/note="T517.2" complement(10039..11244) /gene="At2g39720"
	/note="unknown protein"
	/codon_start=1 /protein_id="AAB87121.1"
	/db_xref="GI:2642154"
	/translation="MASGSWYSCSRFWVWSDSISCPDCDGLLEIOEPLDFPSPD STTTTTHQHSPTRRPPSSSSSPSASMAHDSPTPTVTRSRNSPNVYIVRG SAAAPSDVSEGLDASAQMYIDDOTDGLRLPSPMTELLSGSEFDLLIOISIE LNTNRNLRSCEHPASKSAIEALPLEIOTPHLLSDOSHCAVCKNEVLPKASAREM CNIYHPDILPMLAIRNSCPVCRHELPAEDLTDGAALTATVAEEEDSAAELT IMRLPGGFAVGRIPGMRGDMRPVYVYVTEGRLDRLPRLRVAMSGSRGDRGG GSRERGGFGRIMRLFGCFSSGSSIAAANAASGSGSRIRVTRTRRSFMSPTASS SSRRRWLA"
repeat_region	complement(11453..11512) /rpt_family="POLY_A"
mRNA	complement(join(13323..13599,13724..14005,14087..14559, 14651..14736,14836..14941,15033..15349,15834..16055)) /gene="At2g39730"
mRNA	complement(join(13323..13599,13713..14005,14087..14559, 14651..14736,14836..14941,15033..15349,15834..16055)) /gene="At2g39730"
gene	complement(13323..16055) /gene="At2g39730"
	/note="T517.3"
repeat_region	13364..13461 /rpt_family="POLY_A"
CDS	complement(join(13484..13599,13724..14005,14087..14559, 14651..14736,14836..14941,15033..15349,15834..15878)) /gene="At2g39730"
	/note="alternative splice form 1 Gp1166834 M86720"
	/codon_start=1 /product="Rubsico activase"
	/protein_id="AAB87122.1"
	/db_xref="GI:2642155"
	/translation="MAAASVGAIRNAPLSLNGSGSAGVASPASFLLKVVYTSRE AOSNKSNGSEFKVLAVEREDKQTDGDRRGATDTSDQODITTKGMVDSVEQAPKT GTHNAVLSSEYVSGRLQYNDLNMMDGFIYAPAFMKLVNHTKNEFLPNKVPIL LGIMGCGQCKSFQCELVMAKGINPIMMSAGLESGNAGEPAKLIJORYREAADLIK KGKMCCLFINDLDAGAGMGCTQYTNVNMVATLMNTADNPTNVLQOLGYKKEEA RVPTICTGDNFSTLYAPLIRDGMRKFFYNAPTREDDTCKCKGIFPRDKIKDEDIYTLV LDEPPGQSIDFEGALRAKVIDEVRKFEVSLGVERIKRRLVNSREGPPVEQPMYTEK LMYEGNMLVMEQENVRKVOIAETYLQAAIGDANADAIGGTYGKGAQOQVNPVEG CTDPAVENEDPTARSDDGTCVNF"
CDS	complement(join(13579..13599,13713..14005,14087..14559, 14651..14736,14836..14941,15033..15349,15834..15878)) /gene="At2g39730"
	/note="alternative splice form 2 Gp1166835 M86720"
	/codon_start=1 /product="Rubsico activase"
	/protein_id="AAB87137.1"
	/db_xref="GI:2642170"
	/translation="MAAASVGAIRNAPLSLNGSGSAGVASPASFLLKVVYTSRE AOSNKSNGSEFKVLAVEREDKQTDGDRRGATDTSDQODITTKGMVDSVEQAPKT GTHNAVLSSEYVSGRLQYNDLNMMDGFIYAPAFMKLVNHTKNEFLPNKVPIL LGIMGCGQCKSFQCELVMAKGINPIMMSAGLESGNAGEPAKLIJORYREAADLIK KGKMCCLFINDLDAGAGMGCTQYTNVNMVATLMNTADNPTNVLQOLGYKKEEA RVPTICTGDNFSTLYAPLIRDGMRKFFYNAPTREDDTCKCKGIFPRDKIKDEDIYTLV LDEPPGQSIDFEGALRAKVIDEVRKFEVSLGVERIKRRLVNSREGPPVEQPMYTEK LMYEGNMLVMEQENVRKVOIAETYLQAAIGDANADAIGGTYGKTEKEPEPS"
repeat_region	complement(117561..17639) /rpt_family="(TA)n"
mRNA	join(<18783..18921,19188..19352,19496..19674,19766..19861, 19655..20034,20131..20261,20593..>21237) /gene="At2g39740"
gene	<18783..>21237 /gene="At2g39740"
	/note="T517.4: predicted by genscan"
CDS	join(18783..18921,19188..19352,19496..19674,19766..19861, 19655..20034,20131..20261,20593..21237)

```
/gene="At2g39740"
/!note="hypothetical protein"
/codon_start=1
/protein_id="AAB87123.1"
/db_xref="GI:2642156"
/translation="MSRNPFLDPTLOETILOVLIKPRADRRITIVIDLRLVQSGEC
LRGAVQPFQSVSNLFTRMGDLDISVDFSGSILFTGKKOKQKLLHLALRASG
LMYKLOFVHARVPLIKVSGHORISIDNLDGLKSRFLPSEIDGFEFLVL
LVKEWAKAHINDSKTGPNSSLSLVIHFQRCVPAIPLRLVITPKSAVDLITGV
RKTAEESIAOVTANIANIARFKSRASVNRSSISELVSFEKVEPEEQPVAAASVS
RNRDRIAQVFOITSRLVSECNRSIIIGLTHGHIOESLRTISLPSQIHANGHNV
RNLHGQARPQNDQMOQNSQSYNTPNPMPPLTSPROQNTQNNPRLDGOQPVQO
OTWPIOTQTOQKSPYKSGNRLKNTSAGSSQNGHIGKPSGHNMGVNSAPAYTNG
VNSARPPSKIPSGOIVMRPRHEQ"
complement(join(21380..21726,21839..21916,22124..22246,
22328..22585,22834..22893,23212..23322,23481..23751,
23847..23996,24089..>24814))
/gene="At2g39750"
complement(21380..>24814)
```

Query Match
Best Local Similarity 85.5%; Score 18.8; DB 8; Length 72590;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 tgttgatcacagaattgata 22
||||| ||||||| |||||
Db 35807 TGTGTGTCACAAGATTGTGACA 35828

RESULT 14
AF165146 121524 bp DNA HTG 07-JUN-2001
LOCUS Homo sapiens chromosome 8 clone CTA-397H3 map Bq12-8q13, ***
DEFINITION SEQUENCING IN PROGRESS ***, 7 unordered pieces.
ACCESSION AF165146
VERSION AF165146.3 GI:14327840
KEYWORDS HTG; HTGS_PHASE1; HTGS_FULUTOP.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 121524)
Schlilabel,M.B., Baumgart,C., Blechschmidt,K., Dette,M., Jahn,N.,
Lehmann,R., Menzel,U., Polley,A., Reichwald,K., Schudy,A.,
Siddiqui,R., Taudien,S., Wen,G., Schlegelberger,B., Siebert,R.,
Rosenthal,A. and Platzer,M.
Chromosome 8 genomic sequence
Unpublished
2 (bases 1 to 121524)
Schudy,A., Blechschmidt,K., Schlilabel,M., Baumgart,C., Menzel,U.,
Weber,J., Schattevojt,R. and Rosenthal,A.
Direct Submission
Submitted (06-JUN-1999) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
On Jun 7, 2001 this sequence version replaced g1:8151945.
***** Genome Center
Center: Institute of Molecular Biotechnology
Center code: IMB
Web site: http://genome.imb-jena.de/
Contact: gscj-submit@genome.imb-jena.de
----- Project Information
Center project name: H171
Center clone name: CTA-397H3
----- Summary Statistics
Sequencing vector: M13; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 115981 bases at least Q40
Consensus quality: 117981 bases at least Q30
Consensus quality: 119332 bases at least Q20
Quality coverage: 9.57 in Q20 bases; sum-of-contigs

***** NOTE: This is a 'working draft' sequence. It currently

* consists of 7 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 1187: contig of 1187 bp in length
* 1188 1287: gap of unknown length
* 1288 3530: contig of 2243 bp in length
* 3531 3630: gap of unknown length
* 3631 21624: contig of 17994 bp in length
* 21625 21724: gap of unknown length
* 21725 43656: contig of 21932 bp in length
* 43657 43755: gap of unknown length
* 43757 61383: contig of 17627 bp in length
* 61384 61483: gap of unknown length
* 61484 85318: contig of 23835 bp in length
* 85319 85418: gap of unknown length
* 85419 121524: contig of 36106 bp in length.
Location/Qualifiers
1..121524
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/map="8q12-8q13"
/clone="CTA-397H3"
BASE COUNT 34959 a 25555 c 25579 g 34831 t 600 others
ORIGIN

Query Match
Best Local Similarity 85.5%; Score 18.8; DB 2; Length 121524;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 tgttgatcacagaattgata 22
||||| ||||||| |||||
Db 7103 TTTTGCATCACAAGATTGTGTA 7124

RESULT 15
AC083961
LOCUS Homo sapiens chromosome 8 clone RP11-22E14 map 8, *** SEQUENCING IN
DEFINITION PROGRESS ***, 67 unordered pieces.
ACCESSION AC083961
VERSION AC083961.2 GI:12831382
KEYWORDS HTG; HTGS_PHASE1.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota: Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 129090)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Unpublished
2 (bases 1 to 129090)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Beta,F., Boguslavsky,L.,
Boukigaiter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
Choepel,Y., Colangelo,M., Collins,S., Collamore,A., Cooke,P.,
Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,
Fitzhugh,W., Gage,D., Galagan,J., Gardys,S., Glade,S., Goyette,M.,
Graham,L., Grand-pierre,N., Hagos,B., Heaford,A., Horton,L.,
Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., LaRocque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G.,
Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McKernan,K.,
McPheeters,R., Meldrim,J., Meneses,L., Mhova,T., Mlenga,V.,
Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T.,
O'Donnell,P., O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K.,
Pierre,N., Pisanl,C., Polara,V., Raymond,C., Riebeck,M., Riley,R.,
Rogov,P., Rotman,D., Roy,A., Santos,R., Schauer,S., Severy,P.,
Sougnuez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,

TITLE
JOURNAL
COMMENT

Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Tirrell, A., Travers, M., Tirigillo, J., Vassiliev, H., Viel, R., Vo, A.,
Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zahoun, J.,
Zimmer, A., and Zody, M.
Direct Submission
Submitted (08-OCT-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Feb 15, 2001 this sequence version replaced gi:10717228.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center Project name: L11257

Center Clone name: 22_E_14

* NOTE: This is a 'working draft' sequence. It currently
* consists of 67 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1369: contig of 1369 bp in length
* 1370 1469: gap of 100 bp
* 1470 2279: contig of 810 bp in length
* 2280 2379: gap of 100 bp
* 2380 3192: contig of 813 bp in length
* 3193 3292: gap of 100 bp
* 3293 3539: contig of 247 bp in length
* 3540 3639: gap of 100 bp
* 3640 4530: contig of 891 bp in length
* 4531 4630: gap of 100 bp
* 4631 5236: contig of 606 bp in length
* 5237 5336: gap of 100 bp
* 5337 6055: contig of 719 bp in length
* 6056 6155: gap of 100 bp
* 6156 6653: contig of 498 bp in length
* 6654 6753: gap of 100 bp
* 6754 7388: contig of 635 bp in length
* 7389 7488: gap of 100 bp
* 7489 8145: contig of 657 bp in length
* 8146 8245: gap of 100 bp
* 8246 9151: contig of 906 bp in length
* 9152 9251: gap of 100 bp
* 9252 9962: contig of 711 bp in length
* 9963 10062: gap of 100 bp
* 10063 10776: contig of 714 bp in length
* 10777 10876: gap of 100 bp
* 10877 11474: contig of 598 bp in length
* 11475 11574: gap of 100 bp
* 11575 12639: contig of 1065 bp in length
* 12640 12739: gap of 100 bp
* 12740 13952: contig of 1213 bp in length
* 13953 14052: gap of 100 bp
* 14053 14838: contig of 786 bp in length
* 14839 14938: gap of 100 bp
* 14939 15849: contig of 911 bp in length
* 15850 15949: gap of 100 bp
* 15950 18036: contig of 2087 bp in length
* 18037 18136: gap of 100 bp
* 18137 19282: contig of 1146 bp in length
* 19283 19382: gap of 100 bp
* 19383 20437: contig of 1055 bp in length
* 20438 20537: gap of 100 bp
* 20538 21159: contig of 622 bp in length
* 21160 21259: gap of 100 bp
* 21260 22494: contig of 1235 bp in length

22495 22594: gap of 100 bp
* 22595 23478: contig of 884 bp in length
* 23479 23578: gap of 100 bp
* 23579 24600: contig of 1022 bp in length
* 24601 24700: gap of 100 bp
* 24701 24891: contig of 191 bp in length
* 24892 24991: gap of 100 bp
* 24992 26183: contig of 1192 bp in length
* 26184 26283: gap of 100 bp
* 26284 27862: contig of 1579 bp in length
* 27863 27962: gap of 100 bp
* 27963 28581: contig of 619 bp in length
* 28582 28681: gap of 100 bp
* 28682 29511: contig of 830 bp in length
* 29512 29611: gap of 100 bp
* 29612 31038: contig of 1427 bp in length
* 31039 31138: gap of 100 bp
* 31139 32582: contig of 1444 bp in length
* 32583 32682: gap of 100 bp
* 32683 33540: contig of 858 bp in length
* 33541 33640: gap of 100 bp
* 33641 35263: contig of 1623 bp in length
* 35264 35363: gap of 100 bp
* 35364 37036: contig of 1673 bp in length
* 37037 37136: gap of 100 bp
* 37137 38963: contig of 1827 bp in length
* 38964 39063: gap of 100 bp
* 39064 40807: contig of 1744 bp in length
* 40808 40907: gap of 100 bp
* 40908 42896: contig of 1989 bp in length
* 42897 42996: gap of 100 bp
* 42997 44598: contig of 1602 bp in length
* 44599 44698: gap of 100 bp
* 44699 46343: contig of 1645 bp in length
* 46344 46443: gap of 100 bp
* 46444 47574: contig of 1131 bp in length
* 47575 47674: gap of 100 bp
* 47675 50217: contig of 2543 bp in length
* 50218 50317: gap of 100 bp
* 50318 51338: contig of 1021 bp in length
* 51339 51438: gap of 100 bp
* 51439 53743: contig of 2305 bp in length
* 53744 53843: gap of 100 bp
* 53844 55260: contig of 1417 bp in length
* 55261 55360: gap of 100 bp
* 55361 57843: contig of 2483 bp in length
* 57844 57943: gap of 100 bp
* 57944 59865: contig of 1922 bp in length
* 59866 59965: gap of 100 bp
* 59966 62260: contig of 2295 bp in length
* 62261 62360: gap of 100 bp
* 62361 64120: contig of 1760 bp in length
* 64121 64220: gap of 100 bp
* 64221 65979: contig of 1759 bp in length
* 65980 66079: gap of 100 bp
* 66080 68712: contig of 2633 bp in length
* 68713 68812: gap of 100 bp
* 68813 71145: contig of 2333 bp in length
* 71146 71245: gap of 100 bp
* 71246 74155: contig of 2910 bp in length
* 74156 74255: gap of 100 bp
* 74256 77041: contig of 2786 bp in length
* 77042 77141: gap of 100 bp
* 77142 80685: contig of 3544 bp in length
* 80686 80785: gap of 100 bp
* 80786 83825: contig of 3040 bp in length
* 83826 83925: gap of 100 bp
* 83926 87721: contig of 3796 bp in length
* 87722 87821: gap of 100 bp
* 87822 92728: contig of 4907 bp in length
* 92729 92828: gap of 100 bp
* 92829 96652: contig of 3824 bp in length
* 96653 96752: gap of 100 bp

```

* 96753 99906: contig of 3154 bp in length
* 99907 100006: gap of 100 bp
* 100007 102910: contig of 2904 bp in length
* 102911 103010: gap of 100 bp
* 103011 107188: contig of 4178 bp in length
* 107189 107288: gap of 100 bp
* 107289 111047: contig of 3759 bp in length
* 111048 111147: gap of 100 bp
* 111148 116741: contig of 5594 bp in length
* 116742 116841: gap of 100 bp
* 116842 121990: contig of 5149 bp in length
* 121991 122090: gap of 100 bp
* 122091 126622: contig of 4532 bp in length
* 126623 126722: gap of 100 bp
* 126723 129090: contig of 2368 bp in length.

```

FEATURES

```

source
1. 129090
   /organism="Homo sapiens"
   /db_xref="taxon:9606"
   /chromosome="8"
   /map="8"
   /clone="RP11-22E14"
   /clone_lib="RPC1-11 Human Male BAC"

```

```

Query Match      85.5%; Score 18.8; DB 2; Length 129090;
Best Local Similarity 90.9%; Pred. No. 71;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 tgttgatcacagaatttgata 22
    | ||||| ||||| ||
Db 6294 TTTGGATCACAAGATTGGTA 6315

```

Search completed: December 15, 2001, 03:19:51
 Job time: 7681 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 03:29:41 : Search time 401.91 Seconds
(without alignments)
46.929 Million cell updates/sec

Title: US-09-380-826A-2

Perfect score: 22

Sequence: 1 tttgtgacacagatttgata 22

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : N.Geneseq_1101:*

1: /SIDs2/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SIDs2/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SIDs2/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SIDs2/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SIDs2/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SIDs2/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SIDs2/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SIDs2/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SIDs2/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SIDs2/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SIDs2/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SIDs2/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SIDs2/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SIDs2/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SIDs2/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SIDs2/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SIDs2/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SIDs2/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SIDs2/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SIDs2/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SIDs2/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SIDs2/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	100.0	22	19	AAV58901
2	22	100.0	22	19	AAV58897
3	22	100.0	22	19	AAV58896
4	18.8	85.5	1012	21	AAV99365
5	18.8	85.5	1069	21	AAV37325
6	17.8	80.9	2084	21	AAV9696
7	17.2	78.2	447	21	AAV18114
8	17.2	78.2	1290	21	AAV42365
9	17.2	78.2	1413	21	AAV46258
10	17.2	78.2	2038	22	AAV22799
11	16.8	76.4	1764	22	AAV15352

C	12	16.8	76.4	2582	18	AAV90543	Maize male fertili
C	13	16.4	74.5	2444	21	AAV56059	Eucalyptus grandis
C	14	16.2	73.6	335	22	AAV33702	Human colon cancer
C	15	16.2	73.6	931	20	AAV84598	Human secreted pro
C	16	16.2	73.6	1250	21	AAV59054	Human secreted pro
C	17	16.2	73.6	1291	19	AAV59636	Human secreted pro
C	18	16.2	73.6	1376	22	AAV33007	Human colon cancer
C	19	16.2	73.6	1450	21	AAV49931	Arabidopsis thalia
C	20	16.2	73.6	1452	21	AAV39657	Arabidopsis thalia
C	21	16.2	73.6	1473	20	AAV61757	B. burgdorferi ant
C	22	16.2	73.6	2000	17	AAV17111	Acetyl-CoA-carboxy
C	23	16.2	73.6	2001	17	AAV39904	Maize acetyl CoA c
C	24	16.2	73.6	2001	17	AAV49816	ECORI fragment of
C	25	16.2	73.6	3822	18	AAV46483	M. leprae gyrA pre
C	26	16.2	73.6	4366	14	AAV42933	A3 maize ACCase CD
C	27	16.2	73.6	5400	21	AAV39732	Potato subclone PG
C	28	16.2	73.6	7470	17	AAV39905	Maize acetyl CoA c
C	29	16.2	73.6	7470	19	AAV29317	Maize ACCase enzym
C	30	16.2	73.6	7470	21	AAV49820	Maize acetyl CoA c
C	31	16.2	73.6	10811	19	AAV59091	Potato pollen cell
C	32	15.8	71.8	467	22	AAV32060	Human olfactory re
C	33	15.8	71.8	670	21	AAV09187	Human secreted pro
C	34	15.8	71.8	736	21	AAV10808	Human secreted pro
C	35	15.8	71.8	740	21	AAV51908	Arabidopsis thalia
C	36	15.8	71.8	843	20	AAV15467	Human gene express
C	37	15.8	71.8	961	19	AAV14095	H. pylori GHP0 131
C	38	15.8	71.8	4146	21	AAV25156	Human hypoxia resp
C	39	15.6	70.9	297	20	AAV88753	EST clone HK650.
C	40	15.6	70.9	342	22	AAV32175	Human olfactory re
C	41	15.6	70.9	452	22	AAV5351	S. epidermidis ope
C	42	15.6	70.9	635	22	AAV31988	Human olfactory re
C	43	15.6	70.9	795	20	AAV72025	Adenovirus pACT562
C	44	15.6	70.9	834	20	AAV72026	Adenovirus SCAR-RG
C	45	15.6	70.9	1075	18	AAV72715	C. elegans inhibit

ALIGNMENTS

RESULT	1
ID	AAV58901 standard; DNA: 22 BP.
XX	AAV58901;
AC	AAV58901;
XX	20-JAN-1999 (first entry)
DE	Leptospira rRNA gene nucleotide sequence.
XX	Infection; pathogenic Leptospira; protective immunity; therapy;
KW	diagnosis; ss.
XX	Leptospira sp.
OS	Leptospira sp.
XX	WO9840099-A1.
PN	17-SEP-1998.
XX	06-MAR-1998; 98WO-AU00145.
PF	07-MAR-1997; 97AU-0005494.
XX	(AGRI-) AGRIC VICTORIA SERVICES PTY LTD.
PA	(PIGR-) PIG RES & DEV CORP.
XX	Chappel RJ;
DR	WPI; 1998-520791/44.
XX	New isolated pathogenic Leptospira bacterium - useful for e.g
PT	developing products for conferring protective immunity, and for
XX	prophylactic or therapeutic treatment

PS Claim 15; Page 72; 94pp; English.
XX
CC This sequence represents a Leptospira DNA sequence isolated from the
CC pathogenic Leptospira (LS) bacterium of the invention. The bacterium
CC belongs to serogroup Hurstbridge or serovar hurstbridge or the species
CC L. fainei. The LS bacteria can be used for conferring protective
CC immunity against pathogenic LS bacteria in humans or animals. The
CC bacteria can also be used for prophylactic or therapeutic treatment of LS
CC infections. The DNAs and antibodies may also be used for detection and
CC diagnosis of past or present LS infection.
XX
SQ Sequence 22 BP; 7 A; 2 C; 5 G; 8 T; 0 other;

Query Match 100.0%; Score 22; DB 19; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.11; Mismatches 0; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttgatcacagaatttgata 22
|||||
DB 1 tgttgatcacagaatttgata 22

RESULT 2
AAV58897
ID AAV58897 standard; DNA; 22 BP.
XX
AC AAV58897;
XX
DT 20-JAN-1999 (first entry);
XX
DE L. fainei nucleotide sequence.
XX
KW Infection; pathogenic Leptospira; protective immunity; therapy;
KM diagnosis; ss.
XX
OS Leptospira fainei.
XX
PN WO9840099-A1.
XX
PD 17-SEP-1998.
XX
PF 06-MAR-1998; 98WO-AU00145.
XX
PR 07-MAR-1997; 97AU-0005494.
XX
PA (AGRI-) AGRIC VICTORIA SERVICES PTY LTD.
XX (PIGR-) PIG RES & DEV CORP.
XX
PI Chappel RJ;
XX
DR WPI: 1998-520791/44.
XX
PT New isolated pathogenic Leptospira bacterium - useful for, e.g
PT developing products for conferring protective immunity, and for
PT prophylactic or therapeutic treatment
XX
PS Claim 15; Page 70; 94pp; English.
XX
CC This sequence represents a Leptospira DNA sequence isolated from the
CC pathogenic Leptospira (LS) bacterium of the invention. The bacterium
CC belongs to serogroup Hurstbridge or serovar hurstbridge or the species
CC L. fainei. The LS bacteria can be used for conferring protective
CC immunity against pathogenic LS bacteria in humans or animals. The
CC bacteria can also be used for prophylactic or therapeutic treatment of LS
CC infections. The DNAs and antibodies may also be used for detection and
CC diagnosis of past or present LS infection.
XX
SQ Sequence 22 BP; 7 A; 2 C; 5 G; 8 T; 0 other;

Query Match 100.0%; Score 22; DB 19; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttgatcacagaatttgata 22
|||||
DB 1 tgttgatcacagaatttgata 22

RESULT 3
AAV58896
ID AAV58896 standard; DNA; 1477 BP.
XX
AC AAV58896;
XX
DT 20-JAN-1999 (first entry)
XX
DE L. fainei nucleotide sequence.
XX
KW Infection; pathogenic Leptospira; protective immunity; therapy;
KM diagnosis; ss.
XX
OS Leptospira fainei.
XX
PN WO9840099-A1.
XX
PD 17-SEP-1998.
XX
PF 06-MAR-1998; 98WO-AU00145.
XX
PR 07-MAR-1997; 97AU-0005494.
XX
PA (AGRI-) AGRIC VICTORIA SERVICES PTY LTD.
XX (PIGR-) PIG RES & DEV CORP.
XX
PI Chappel RJ;
XX
DR WPI: 1998-520791/44.
XX
PT New isolated pathogenic Leptospira bacterium - useful for, e.g
PT developing products for conferring protective immunity, and for
PT prophylactic or therapeutic treatment
XX
PS Claim 15; Page 69-70; 94pp; English.
XX
CC This sequence represents a Leptospira DNA sequence isolated from the
CC pathogenic Leptospira (LS) bacterium of the invention. The bacterium
CC belongs to serogroup Hurstbridge or serovar hurstbridge or the species
CC L. fainei. The LS bacteria can be used for conferring protective
CC immunity against pathogenic LS bacteria in humans or animals. The
CC bacteria can also be used for prophylactic or therapeutic treatment of LS
CC infections. The DNAs and antibodies may also be used for detection and
CC diagnosis of past or present LS infection.
XX
SQ Sequence 1477 BP; 390 A; 334 C; 439 G; 314 T; 0 other;

Query Match 100.0%; Score 22; DB 19; Length 1477;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttgatcacagaatttgata 22
|||||
DB 154 tgttgatcacagaatttgata 175

RESULT 4
AAA99365
ID AAA99365 standard; DNA; 1012 BP.
XX
AC AAA99365;
XX
DT 22-JAN-2001 (first entry)
XX
DE Plant PRG1 promoter related gene sequence.

XX Plant promoter; PrAG1; reproductive tissue; transgenic plant; cereal; ds.
XX Arabidopsis thaliana.
OS WO200055172-A1.
XX 21-SEP-2000.
PD 17-MAR-2000; 2000WO-NZ00031.
XX 17-MAR-1999; 99NZ-0334715.
PR (CART-) CARTER HOLT HARVEY LTD.
PA (TASM-) TASMAN BIOTECHNOLOGY LTD.
PA (UNMF) UNIV MICHIGAN TECHNOLOGICAL.
XX Podila GK, Liu J, Karnosky DF;
PI WPI: 2000-594442/56.
DR P-PSDB: AAB26796.
XX Novel plant reproductive tissue promoter, useful to produce plants
PT which have a diminished reproductive capacity or which are sterile
XX Claim 15; Page 42-43; 51pp; English.
PS This invention relates to a novel plant promoter gene. The promoter is
XX located in plant reproductive tissue, and the invention includes
CC transgenic plants containing the promoter. The promoter can be used to
CC produce plants which have a diminished reproductive capacity or which are
CC sterile. The constructs can also be used to transform agronomically
CC important plants in which modulation of reproductive capacity
CC (particularly the timing and abundance of flowering) is desirable,
CC e.g. cereals, rice, maize, wheat, barley, oats, rye, soybean and canola.
CC The present sequence represents DNA encoding the plant reproductive
CC promoter (PrAG1) of the invention.
XX Sequence 1012 BP; 287 A; 197 C; 217 G; 311 T; 0 other;
SO

Query Match 85.5%; Score 18.8; DB 21; Length 1012;
Best Local Similarity 90.9%; Pred. No. 5.8;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ttttgatcacaaagtattgata 22
DB 672 tgttgggtcacaaagtattgaca 693

RESULT 5
AAC37325
ID AAC37325 standard; DNA; 1069 BP.
XX AAC37325;
AC 17-OCT-2000 (first entry)
DT Arabidopsis thaliana DNA fragment SEQ ID NO: 16971.
DE Arabidopsis thaliana
XX Hybridisation assay; genetic mapping; gene expression control;
XX protein identification; signal transduction pathway;
KM metabolic pathway; promoter; termination sequence; ss.
XX Arabidopsis thaliana.
OS Arabidopsis thaliana.
XX EP1033405-A2.
PN 06-SEP-2000.
PD 25-FEB-2000; 2000EP-0301439.
PF 25-FEB-1999; 99US-0121825.1
XX 25-FEB-1999; 99US-0121825.1
PR

PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135153.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 28-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139859.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.

DR WPI: 2000-476052/41.
XX Isolated polynucleotide encoding a polypeptide involved in cell
PT signaling used for generating transgenic plants with modified responses
XX to external signals -
PS Claim 1: Page 408-409; 527pp; English.
XX
CC AA79263 to AA79736 and AAB25100 to AAB2570 represent polynucleotide
CC and protein sequences isolated from eucalyptus (Eucalyptus grandis) or
CC pine (Pinus radiata also known as Monterey pine). The protein sequences
CC are involved in cell signaling. The polynucleotide and protein
CC sequences can be used to modify the response of plant cells to external
CC signals e.g. environmental changes or pathogens during the growth and
CC development of a plant. They can be used to modify cell proliferation,
CC differentiation, elongation and survival, resistance to disease and
CC nutrient metabolism. Examples of modifications which can be produced are
CC altered fruit ripening and senescence of leaves and flowers e.g. to
CC delay senescence and prolong the life of cut flowers or enhance
CC senescence of reproductive organs to engineer sterile plants. Other
CC modifications can be used to delay senescence in selected cell types or
CC organs providing fruit and vegetables which have a longer shelf life
CC between harvest and consumption, or to decrease branching frequency in
CC forest tree species giving long stretches of valuable knot-free clear
CC wood which can be used in solid timber furniture and veneers.
XX
SQ Sequence 2084 BP; 502 A; 434 C; 566 G; 582 T; 0 other;

Query Match 80.9%; Score 17.8; DB 21; Length 2084;
Best Local Similarity 90.5%; Pred. No. 19;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ttttgatcacagaattgac-21
||||| ||||| |||||
DB 1347 TGTGTGACGACCAAGTTTGAT 1327.

RESULT 7
AAC18114
ID AAC18114 standard: cDNA; 447 BP.
XX
AC AAC18114;
XX
DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein 5' EST, SEQ ID NO: 22189;
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX
OS Homo sapiens.
XX
PN EP1033401-A2.
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-0200610.
XX
PR 26-FEB-1999; 990S-0122487.
XX
PA (GEST) GENSET.
XX
PI Dumas Mline Edwards J, Duclert A, Giordano J;
XX
DR WPI: 2000-500381/45.
XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
PS Claim 1; SEQ ID 22189; 71pp + CD-ROM; English.
XX

CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.
XX
SQ Sequence 447 BP; 187 A; 57 C; 66 G; 128 T; 9 other;

Query Match 78.2%; Score 17.2; DB 21; Length 447;
Best Local Similarity 86.4%; Pred. No. 32;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ttttgatcacagaattgata 22
||||| ||||| || |||
DB 240 ttttgatcacacaactaata 261

RESULT 8
AAC42365/C
ID AAC42365 standard: DNA; 1290 BP.
XX
AC AAC42365;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 35275.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 990S-0121825.
PR 05-MAR-1999; 990S-0123180.
PR 09-MAR-1999; 990S-0123548.
PR 23-MAR-1999; 990S-0125788.
PR 25-MAR-1999; 990S-0126264.
PR 29-MAR-1999; 990S-0126785.
PR 01-APR-1999; 990S-0127462.
PR 06-APR-1999; 990S-0128234.
PR 08-APR-1999; 990S-0128714.
PR 16-APR-1999; 990S-0129645.
PR 19-APR-1999; 990S-0130077.
PR 21-APR-1999; 990S-0130449.
PR 23-APR-1999; 990S-0130510.
PR 28-APR-1999; 990S-0130891.
PR 30-APR-1999; 990S-0131449.
PR 30-APR-1999; 990S-0132048.
PR 30-APR-1999; 990S-0132407.
PR 04-MAY-1999; 990S-0132484.
PR 05-MAY-1999; 990S-0132485.
PR 06-MAY-1999; 990S-0132486.
PR 07-MAY-1999; 990S-0132487.
PR 11-MAY-1999; 990S-0132863.
PR 14-MAY-1999; 990S-0134256.
PR 990S-0134218.

PR 14-MAY-1999; 9905-0134219.
PR 14-MAY-1999; 9905-0134221.
PR 14-MAY-1999; 9905-0134370.
PR 18-MAY-1999; 9905-0134768.
PR 19-MAY-1999; 9905-0134941.
PR 20-MAY-1999; 9905-0135124.
PR 21-MAY-1999; 9905-0135353.
PR 24-MAY-1999; 9905-0135629.
PR 25-MAY-1999; 9905-0136021.
PR 27-MAY-1999; 9905-0136392.
PR 28-MAY-1999; 9905-0136782.
PR 01-JUN-1999; 9905-0137222.
PR 03-JUN-1999; 9905-0137528.
PR 04-JUN-1999; 9905-0137502.
PR 07-JUN-1999; 9905-0137724.
PR 08-JUN-1999; 9905-0138094.
PR 10-JUN-1999; 9905-0138540.
PR 10-JUN-1999; 9905-0138847.
PR 14-JUN-1999; 9905-0139119.
PR 16-JUN-1999; 9905-0139452.
PR 16-JUN-1999; 9905-0139453.
PR 17-JUN-1999; 9905-0139492.
PR 18-JUN-1999; 9905-0139454.
PR 18-JUN-1999; 9905-0139455.
PR 18-JUN-1999; 9905-0139456.
PR 18-JUN-1999; 9905-0139457.
PR 18-JUN-1999; 9905-0139458.
PR 18-JUN-1999; 9905-0139459.
PR 18-JUN-1999; 9905-0139460.
PR 18-JUN-1999; 9905-0139461.
PR 18-JUN-1999; 9905-0139462.
PR 18-JUN-1999; 9905-0139463.
PR 18-JUN-1999; 9905-0139750.
PR 18-JUN-1999; 9905-0139763.
PR 21-JUN-1999; 9905-0139817.
PR 22-JUN-1999; 9905-0139899.
PR 23-JUN-1999; 9905-0140353.
PR 23-JUN-1999; 9905-0140354.
PR 24-JUN-1999; 9905-0140685.
PR 28-JUN-1999; 9905-0140828.
PR 29-JUN-1999; 9905-0140991.
PR 30-JUN-1999; 9905-0141287.
PR 01-JUL-1999; 9905-0141642.
PR 01-JUL-1999; 9905-0142154.
PR 02-JUL-1999; 9905-0142055.
PR 06-JUL-1999; 9905-0142390.
PR 08-JUL-1999; 9905-0142803.
PR 09-JUL-1999; 9905-0142920.
PR 12-JUL-1999; 9905-0142977.
PR 13-JUL-1999; 9905-0143542.
PR 14-JUL-1999; 9905-0143624.
PR 15-JUL-1999; 9905-0144005.
PR 16-JUL-1999; 9905-0144085.
PR 16-JUL-1999; 9905-0144086.
PR 19-JUL-1999; 9905-0144325.
PR 19-JUL-1999; 9905-0144331.
PR 19-JUL-1999; 9905-0144332.
PR 19-JUL-1999; 9905-0144333.
PR 19-JUL-1999; 9905-0144334.
PR 19-JUL-1999; 9905-0144335.
PR 20-JUL-1999; 9905-0144352.
PR 20-JUL-1999; 9905-0144632.
PR 20-JUL-1999; 9905-0144884.
PR 21-JUL-1999; 9905-0144814.
PR 21-JUL-1999; 9905-0145086.
PR 21-JUL-1999; 9905-0145088.
PR 22-JUL-1999; 9905-0145085.
PR 22-JUL-1999; 9905-0145087.
PR 22-JUL-1999; 9905-0145089.
PR 22-JUL-1999; 9905-0145192.
PR 23-JUL-1999; 9905-0145145.
PR 23-JUL-1999; 9905-0145218.
PR 23-JUL-1999; 9905-0145224.

PR 26-JUL-1999; 9905-0145276.
PR 27-JUL-1999; 9905-0145913.
PR 27-JUL-1999; 9905-0145918.
PR 27-JUL-1999; 9905-0145919.
PR 28-JUL-1999; 9905-0145951.
PR 02-AUG-1999; 9905-0146386.
PR 02-AUG-1999; 9905-0146388.
PR 02-AUG-1999; 9905-0146389.
PR 03-AUG-1999; 9905-0147038.
PR 04-AUG-1999; 9905-0147204.
PR 04-AUG-1999; 9905-0147302.
PR 05-AUG-1999; 9905-0147192.
PR 05-AUG-1999; 9905-0147260.
PR 06-AUG-1999; 9905-0147303.
PR 06-AUG-1999; 9905-0147416.
PR 09-AUG-1999; 9905-0147493.
PR 09-AUG-1999; 9905-0147935.
PR 10-AUG-1999; 9905-0148171.
PR 11-AUG-1999; 9905-0148319.
PR 12-AUG-1999; 9905-0148341.
PR 13-AUG-1999; 9905-0148565.
PR 13-AUG-1999; 9905-0148684.
PR 15-AUG-1999; 9905-0149368.
PR 17-AUG-1999; 9905-0149175.
PR 18-AUG-1999; 9905-0149426.
PR 20-AUG-1999; 9905-0149722.
PR 20-AUG-1999; 9905-0149723.
PR 20-AUG-1999; 9905-0149929.
PR 23-AUG-1999; 9905-0149902.
PR 23-AUG-1999; 9905-0149930.
PR 25-AUG-1999; 9905-0150586.
PR 26-AUG-1999; 9905-0150884.
PR 27-AUG-1999; 9905-0151065.
PR 27-AUG-1999; 9905-0151066.
PR 27-AUG-1999; 9905-0151080.
PR 30-AUG-1999; 9905-0151303.
PR 31-AUG-1999; 9905-0151438.
PR 01-SEP-1999; 9905-0151930.
PR 07-SEP-1999; 9905-0152363.
PR 10-SEP-1999; 9905-0153070.
PR 13-SEP-1999; 9905-0153758.
PR 15-SEP-1999; 9905-0154018.
PR 16-SEP-1999; 9905-0154039.
PR 20-SEP-1999; 9905-0154779.
PR 22-SEP-1999; 9905-0155139.
PR 23-SEP-1999; 9905-0155486.
PR 24-SEP-1999; 9905-0155659.
PR 28-SEP-1999; 9905-0156458.
PR 29-SEP-1999; 9905-0156586.
PR 04-OCT-1999; 9905-0157117.
PR 05-OCT-1999; 9905-0157753.
PR 06-OCT-1999; 9905-0157865.
PR 07-OCT-1999; 9905-0158029.
PR 08-OCT-1999; 9905-0158232.
PR 12-OCT-1999; 9905-0158369.
PR 13-OCT-1999; 9905-0159293.
PR 13-OCT-1999; 9905-0159294.
PR 13-OCT-1999; 9905-0159295.
PR 14-OCT-1999; 9905-0159329.
PR 14-OCT-1999; 9905-0159330.
PR 14-OCT-1999; 9905-0159331.
PR 14-OCT-1999; 9905-0159637.
PR 14-OCT-1999; 9905-0159638.
PR 18-OCT-1999; 9905-0159584.
PR 21-OCT-1999; 9905-0160741.
PR 21-OCT-1999; 9905-0160767.
PR 21-OCT-1999; 9905-0160768.
PR 21-OCT-1999; 9905-0160770.
PR 21-OCT-1999; 9905-0160814.
PR 21-OCT-1999; 9905-0160815.
PR 22-OCT-1999; 9905-0160980.
PR 22-OCT-1999; 9905-0160981.
PR 22-OCT-1999; 9905-0160989.

```
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 78.2%; Score 17.2; DB 21; Length 1290;
Best Local Similarity 86.4%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 1 ttttgatcacagaattgata 22
Db 642 ttttgcatacacaagatttcttta 621

RESULT 9
AAC46258/c
ID AAC46258 standard; DNA; 1413 BP.
XX
XX AAC46258;
XX
XX 18-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 49483.
DE
XX
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 49483.
XX
XX Hybridization assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
OS
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 16-APR-1999; 99US-0128714.
PR 19-APR-1999; 99US-0129845.
PR 21-APR-1999; 99US-0130077.
PR 23-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146387.
```

```

PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 04-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.

```

```

PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 78.28; Score 17.2; DB 21; Length 1413;
Best Local Similarity 86.4%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ttgttgatcacagaatttgata 22
Db 576 TGTGTGTAACAAGATTGTGA 555

RESULT 10
AAS22799
ID AAS22799 standard; cDNA; 2038 BP.
XX
AC AAS22799;
XX
DT 24-OCT-2001 (first entry)
XX
DE Human cDNA encoding a novel human protein #365.
XX
KW Human; novel protein; ss; Antianaemic; osteopathic; antiinflammatory;
KW Immunomodulatory; cytosolic; neuroprotective; vulnary; nootropic;
KW anticonvulsant; antiarthritic; cerebroprotective; antifungal; antiviral;
KW antibacterial; antiallergic; dermatological; haemostatic; antisthmatic;
KW thrombolytic; immunogen; antibody; gene therapy; neurological disorder;
KW Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;
KW tissue regeneration; immune disorder.
XX
OS Homo sapiens.
XX
PN WO200155437-A2.
XX
PD 02-AUG-2001.
XX
PF 25-JAN-2001; 2001MO-USO2623.
XX
PR 25-JAN-2000; 2000US-0491404.
XX
PA (HYSE-) HYSEO INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
XX WPI: 2001-451939/48.
XX
XX P-PSDB; AAU14494.
XX
PT Isolated polypeptides useful for treating anti-inflammatory diseases,
PT nervous system disorders, and for regenerating bone and cartilage -
XX
XX Claim 1; Page 733-734; 894pp; English.
XX
CC The invention relates to polynucleotides encoding novel human
CC proteins or their active domains. The polypeptides, polynucleotides and
CC antibodies raised against the polypeptides are used in a method of
CC treatment of a mammal and prevention of disorders caused by the aberrant
CC protein expression or activity. The polypeptides can be used as
CC molecular weight markers, food supplements, and in antibody production.
CC The polypeptides are used to identify compounds which bind to the
CC polypeptides. Polynucleotides of the invention are used as probes and
CC primers, for sequencing, for chromosome or gene mapping, in the
CC production of recombinant proteins, and in generating anti-sense DNA or
CC RNA and in gene therapy. Polypeptides of the invention can be used to
CC target drugs to a tumour, in assays to determine biological activity, to
CC raise antibodies/ elicit an immune response, to determine quantitative
CC protein levels, as tissue markers, and to isolate receptors or ligands.
CC Polypeptides of the invention may also be useful in treating platelet
CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,
CC ligament and/or nerve tissue, wound healing, treating burns, promoting
CC the proliferation, differentiation and survival of stem cells, as a
CC contraceptive, treating osteoporosis and osteoarthritis, anaemia,
CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral

```

CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or
CC fungal infection or from autoimmunity, cancer, allergy, asthma,
CC graft-versus-host disease, eczema, haemophilia, thrombosis,
CC anti-inflammatory diseases, nervous system disorders, and infection.
CC The present sequence encodes a protein of the invention.

XX Sequence 2038 BP; 658 A; 400 C; 503 G; 477 T; 0 other;

Query Match 78.2%; Score 17.2; DB 22; Length 2038;

Best Local Similarity 86.4%; Pred. No. 38;

Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 tgttgatcacaaagtattga 22
||||||| ||||||| ||

Db 1929 tgttgagacacaagtattaa 1950

RESULT 11

AAH15352/c
ID AAH15352 standard; cDNA; 1764 BP.

AC AAH15352;

DT 26-JUN-2001 (first entry)

DE Human cDNA sequence SEQ ID NO:13526.

XX Human; primer: detection; diagnosis; antisense therapy; gene therapy; ss.

XX Homo sapiens.

OS EP1074617-A2.

PN 07-FEB-2001.

PD 28-JUL-2000; 2000EP-0116126.

PF 29-JUL-1999; 99JP-0248036.

PR 27-AUG-1999; 99JP-0300253.

PR 11-JAN-2000; 2000JP-0118776.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

PA (HELI-) HELIX RES INST.

PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

DR WPI; 2001-318749/34.

PS Claim 8; SEQ ID 13526; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the

CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.

XX Sequence 1764 BP; 511 A; 347 C; 392 G; 514 T; 0 other;

Query Match 76.4%; Score 16.8; DB 22; Length 1764;

Best Local Similarity 90.0%; Pred. No. 58;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 tgttgatcacaaagtattga 20
||||||| ||||||| ||

Db 783 tgttgatcacaaagtattga 764

RESULT 12

AAT90543/c
ID AAT90543 standard; DNA; 2582 BP.

AC AAT90543;

DT 13-FEB-1998 (first entry)

DE Maize male fertility gene Zm41-A.

XX Zm41-A; male fertility; hybrid seed; transgenic plant; antisense;

XX ribozyme; male sterile; maize; Ms41-A; ds.

XX Zea mays.

OS Key Location/Qualifiers

FT exon 823..915

FT intron 916..935

FT exon 936..1090

FT intron 1091..1167

FT exon 1168..1332

FT exon 1333..1332

FT exon 1333..1332

FT exon 1333..1332

FT exon 1333..1332

FT exon 1333..1332

FT exon 1333..1332

FT exon 1333..1332

FT exon 1333..1332

FT exon 1333..1332

FT exon 1333..1332

FT exon 1333..1332

FT exon 1333..1332

FT exon 1333..1332

FT exon 1333..1332

FT exon 1333..1332

CC be derived from 231 via genetic rearrangements, deletions and/or
CC insertions. 233 has subsequent deletions from 235 and is
CC truncated, having only exons 3, 5 and 6. The MS41-A and Zm41-A
CC genes, antisense or ribozyme sequences can be used to produce
CC transgenic plants with controlled male fertility. Male sterile
CC plants are useful for hybrid seed production, particularly in
CC Brassicaceae, lettuce, spinach and onions.
CC
XX
SQ Sequence 2582 BP; 670 A; 512 C; 537 G; 863 T; 0 other;

Query Match 76.4%; Score 16.8; DB 18; Length 2582;
Best Local Similarity 90.0%; Pred. No. 60;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 tgttgatcaccaagattga 20
Db 268 TGTGTGATTCACAGATCTGA 249
||||| ||||||| |||

RESULT 13
AAC56059
ID AAC56059 standard; DNA: 2444 BP.
AC AAC56059,
XX
XX
DT 25-JAN-2001 (first entry)
XX
XX
DE Eucalyptus grandis transcription factor DNA sequence #190.
XX
XX
KW Plant; transcription factor; gene expression; eucalyptus; pine; acacia;
KW poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor;
KW basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;
KW homeodomain zipper; LIM domain; AP2; EREBs; zinc finger domain;
KW type 2 Cys2His2; CCAAT box element; MYB; ss.
XX
XX
OS Eucalyptus grandis.
XX
XX
PN WO200053724-A2.
XX
PD 14-SEP-2000.
XX
PF 09-MAR-2000; 2000MO-US06112.
XX
PR 11-MAR-1999; 99US-0266513.
PR 18-AUG-1999; 99US-0149485.
XX
XX
PA (GENE-) GENESIS RES & DEV CORP LTD.
PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX
PI Wood M, McGrath A, Shenk MA, Glenn M;
XX
XX
DR WPI: 2000-579369/54.
XX
XX
PT New isolated polynucleotide encoding a plant transcription factor for
PT producing a plant e.g. a woody plant, preferably eucalyptus or pine,
PT having modified gene expression or modified activity of a polypeptide
XX
XX
PS Claim 1; Pages 94-95; 747pp; English.

CC The present invention relates to novel plant transcription factors from
CC Eucalyptus grandis or Pinus radiata. The present sequence is the coding
CC sequence for one such transcription factor. The transcription factor may
CC be used to produce a plant having modified gene expression such as a
CC woody plant e.g. a eucalyptus, pine, acacia, poplar, sweetgum, teak, or
CC mahogany species or to modify the activity of a polypeptide in a plant.
CC The transcription factors of the present invention are members from the
CC following families of regulatory proteins: bZIP, bZIP family of G-box
CC binding factors, basic helix-loop-helix zipper,
CC homeotic/homeodomain/homeobox/MADS, homeodomain zipper, LIM domain, AP2
CC and EREBs, zinc finger domains of type 2 Cys2His2, CCAAT box elements

CC and MYB.
XX
SQ Sequence 2444 BP; 643 A; 554 C; 640 G; 607 T; 0 other;

Query Match 74.5%; Score 16.4; DB 21; Length 2444;
Best Local Similarity 94.4%; Pred. No. 94;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 tgttgatcaccaagattt 18
Db 1451 tgatgacacacagattt 1468
||||| ||||||| |||

RESULT 14
AAH33702/C
ID AAH33702 standard; cDNA; 335 BP.
AC AAH33702;
XX
XX
DT 03-SEP-2001 (first entry)
XX
XX
DE Human colon cancer antigen encoding cDNA SEQ ID NO:758.
XX
XX
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW colorectal carcinoma; ss.
XX
XX
OS Homo sapiens.
XX
XX
PN WO200122920-A2.
XX
XX
PD 05-APR-2001.
XX
XX
PF 28-SEP-2000; 2000MO-US26524.
XX
XX
PR 29-SEP-1999; 99US-0157137.
PR 03-NOV-1999; 99US-0163280.
XX
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX
PI Ruben SM, Barash SC, Birse CE, Rosen CA;
XX
XX
DR WPI: 2001-235357/24.
DR P-PSDB; AAG74271.
XX
XX
PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers -
XX
XX
PS Claim 1; Page 2738; 9803pp; English.

CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patients own production of P.
CC Additionally, N may be used to produce the colon cancer-associated Ps
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAB77789 represent sequences used in the exemplification of the
CC present invention.
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.
XX
SQ Sequence 335 BP; 95 A; 71 C; 60 G; 105 T; 4 other;


```

Query Match          73.6%; Score 16.2; DB 22; Length 335;
Best Local Similarity 85.7%; Pred. No. 96;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy      1 ttttgatcacaaagttagat 21
        ||||| | ||||| ||||| ||
Db      122 TCTGTGACGACAAAGATTAT 102

RESULT 15
AAV84598/C
ID      AAV84598 standard; DNA; 951 BP.
XX
XX      AAV84598;
AC
XX
XX      01-MAR-1999 (first entry)
DT
XX
XX      Human secreted protein gene 188 clone HHP5F70.
DE
XX
XX      Human; secreted protein; fusion protein; gene therapy; protein therapy;
KM      diagnosis; cancer; tumour; neurodegenerative disorder; leukaemia;
KM      developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
KM      immune system; ascites; lymphocytic disease; brain; hepatitis; lymphoma;
KM      inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KM      cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KM      osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KM      endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
XX      Homo sapiens.
OS
XX
XX      WO9854963-A2.
PN
XX
XX      10-DEC-1998.
PD
XX
XX      04-JUN-1998; 98MO-US11422.
PF
XX
XX      18-DEC-1997; 97US-0070923.
PR      06-JUN-1997; 97US-0048877.
PR      06-JUN-1997; 97US-0048881.
PR      06-JUN-1997; 97US-0048884.
PR      06-JUN-1997; 97US-0048893.
PR      06-JUN-1997; 97US-0048896.
PR      06-JUN-1997; 97US-0048899.
PR      06-JUN-1997; 97US-0048915.
PR      06-JUN-1997; 97US-0048949.
PR      06-JUN-1997; 97US-0048964.
PR      06-JUN-1997; 97US-0048972.
PR      06-JUN-1997; 97US-0049020.
PR      06-JUN-1997; 97US-0049375.
PR      05-SEP-1997; 97US-0057628.
PR      05-SEP-1997; 97US-0057635.
PR      05-SEP-1997; 97US-0057647.
PR      05-SEP-1997; 97US-0057647.
PR      05-SEP-1997; 97US-0057660.
PR      05-SEP-1997; 97US-0057661.
PR      05-SEP-1997; 97US-0057667.
PR      05-SEP-1997; 97US-0057761.
PR      05-SEP-1997; 97US-0057764.
PR      05-SEP-1997; 97US-0057770.
PR      05-SEP-1997; 97US-0057775.
PR      05-SEP-1997; 97US-0057778.
PR      06-JUN-1997; 97US-0048875.
PR      06-JUN-1997; 97US-0048878.
PR      06-JUN-1997; 97US-0048882.
PR      06-JUN-1997; 97US-0048885.
PR      06-JUN-1997; 97US-0048894.
PR      06-JUN-1997; 97US-0048897.
PR      06-JUN-1997; 97US-0048900.
PR      06-JUN-1997; 97US-0048916.
PR      06-JUN-1997; 97US-0048962.
PR      06-JUN-1997; 97US-0048970.
PR      06-JUN-1997; 97US-0048974.
PR      06-JUN-1997; 97US-0049373.

```

```

PR      05-SEP-1997; 97US-0057584.
PR      05-SEP-1997; 97US-0057629.
PR      05-SEP-1997; 97US-0057642.
PR      05-SEP-1997; 97US-0057645.
PR      05-SEP-1997; 97US-0057648.
PR      05-SEP-1997; 97US-0057651.
PR      05-SEP-1997; 97US-0057662.
PR      05-SEP-1997; 97US-0057668.
PR      05-SEP-1997; 97US-0057762.
PR      05-SEP-1997; 97US-0057765.
PR      05-SEP-1997; 97US-0057771.
PR      05-SEP-1997; 97US-0057776.
PR      06-JUN-1997; 97US-0048876.
PR      06-JUN-1997; 97US-0048880.
PR      06-JUN-1997; 97US-0048883.
PR      06-JUN-1997; 97US-0048892.
PR      06-JUN-1997; 97US-0048895.
PR      06-JUN-1997; 97US-0048898.
PR      06-JUN-1997; 97US-0048901.
PR      06-JUN-1997; 97US-0048917.
PR      06-JUN-1997; 97US-0048963.
PR      06-JUN-1997; 97US-0048971.
PR      06-JUN-1997; 97US-0049019.
PR      06-JUN-1997; 97US-0049374.
PR      05-SEP-1997; 97US-0057627.
PR      05-SEP-1997; 97US-0057634.
PR      05-SEP-1997; 97US-0057643.
PR      05-SEP-1997; 97US-0057646.
PR      05-SEP-1997; 97US-0057649.
PR      05-SEP-1997; 97US-0057654.
PR      05-SEP-1997; 97US-0057666.
PR      05-SEP-1997; 97US-0057760.
PR      05-SEP-1997; 97US-0057763.
PR      05-SEP-1997; 97US-0057769.
PR      05-SEP-1997; 97US-0057774.
PR      05-SEP-1997; 97US-0057777.
XX
XX      (HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX      Brewer LA, Carter KC, Dillon PJ, Ehnert R, Endress GA;
PI      Fan P, Feng P, Ferlie AM, Fischer CU, Florence CJ;
PI      Florence K, Greene JM, Hu J, Kyaw H, Lafleur DW;
PI      Li Y, Moore PA, Ni J, Olsen HS, Rosen CA, Ruben SM;
PI      Shi Y, Soppet DR, Wei Y, Young P, Yu G, Zeng Z;
XX
XX      MPI: 1999-059865/05.
DR      P-PSDB; AAW88721.
DR
XX
XX      New isolated human genes and the secreted polypeptides they encode -
XX      useful for diagnosis and treatment of e.g. cancers, neurological
XX      disorders, immune diseases, inflammation or blood disorders
XX
XX      Claim 4; Page 450-451; 772pp; English.
XX
XX      The invention relates to nucleic acid sequences (AAV84411 to AAV84633)
XX      encoding human secreted proteins (AAW88534 to AAW88756). The secreted
XX      protein gene sequences are deposited with the ATCC under deposit numbers
XX      ATCC 97979, 97974, 97975, 97976, 97977, 209007, 209008, 209009, 209010,
XX      209011, 209080, 209081, 209082, 209083, 209084, 209085, 209511. Host
XX      cells comprising recombinant vectors containing the nucleic acid
XX      sequences are used for the recombinant production of the secreted
XX      proteins. The polynucleotide and amino acid sequences are useful for are
XX      useful for preventing, treating or ameliorating medical conditions e.g.
XX      by protein or gene therapy. Pathological conditions can be also
XX      diagnosed by determining the amount of the new polypeptides in a sample
XX      or by determining the presence of mutations in the new polynucleotides.
XX      Specific uses are described for each of the polynucleotides, based on
XX      which tissues they are most highly expressed in, and include developing
XX      CC products for the diagnosis or treatment of cancer, neurodegenerative
XX      disorders, developmental abnormalities and foetal deficiencies, blood
XX      disorders, tumours, leukemias, diseases of the immune system, autoimmune
XX      diseases, hepatic and renal disease, lymphomas, inflammation, allergies,
XX      ischemic shock, Alzheimer's and cognitive disorders, schizophrenia,

```

CC restenosis, prostate diseases, obesity, disorders involving osteoclasts
 CC such as osteoporosis, arthritis or malignancies, diseases of testes,
 CC lung or thymus, digestive/endocrine disorders, infections and AIDS. The
 CC polypeptides are also useful for identifying their binding partners.
 CC The present sequence represents a gene encoding a human secreted protein
 CC (see descriptor line for gene number and clone identification).
 CC
 XX

SQ Sequence 951 BP; 296 A; 134 C; 192 G; 329 T; 0 other;

Query Match 73.68; Score 16.2; DB 20; Length 951;
 Best Local Similarly 85.7%; Pred. No. 1.1e+02;

Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ttttgatcacagatttgat 21
 ||||| || ||||| |||||

Db 524 TGTGTGTTCTCAGAGCTCGAT 504

Search completed: December 15, 2001, 03:29:43
 Job time: 7883 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 03:22:44 ; Search time 172.39 Seconds
(without alignments)
28.903 Million cell updates/sec

Title: US-09-380-826A-2

Perfect score: 22

Sequence: 1 tcttgatcacacaagattgata 22

Scoring table: IDENTITY_NUC

Searched: Gapop 10.0 , Gapext 1.0

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/lna/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/lna/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/lna/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/lna/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/lna/PCUTS.COMB.seq:*
6: /cgn2_6/ptodata/2/lna/Backfillseq1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	16.8	76.4	2582	US-09-102-528-31	Sequence 31, Appl
2	16.2	73.6	2000	US-08-014-326-1	Sequence 1, Appl
3	16.2	73.6	2001	US-08-417-089-1	Sequence 1, Appl
4	16.2	73.6	2001	US-08-695-651-1	Sequence 1, Appl
5	16.2	73.6	2001	US-08-930-285-1	Sequence 1, Appl
6	16.2	73.6	2001	US-08-695-421-1	Sequence 1, Appl
7	16.2	73.6	4345	US-08-244-537-1	Sequence 1, Appl
8	16.2	73.6	7470	US-08-417-089-5	Sequence 1, Appl
9	16.2	73.6	7470	US-08-695-651-5	Sequence 5, Appl
10	16.2	73.6	7470	US-08-930-285-5	Sequence 5, Appl
11	16.2	73.6	7470	US-08-695-421-5	Sequence 5, Appl
12	15.6	70.9	296	US-08-602-145-14	Sequence 14, Appl
13	15.6	70.9	1095	US-08-928-383B-3	Sequence 3, Appl
14	15.6	70.9	1096	US-08-684-862-8	Sequence 8, Appl
15	15.6	70.9	1584	US-08-928-383B-1	Sequence 1, Appl
16	15.6	70.9	1989	US-08-792-055-1	Sequence 1, Appl
17	15.6	70.9	2434	US-09-272-496-1	Sequence 1, Appl
18	15.6	70.9	2770	US-09-008-697A-13	Sequence 13, Appl
19	15.6	70.9	7739	US-09-195-966-1	Sequence 1, Appl
20	15.6	70.9	7745	US-08-299-675-1	Sequence 1, Appl
21	15.6	70.9	7745	US-08-485-241-1	Sequence 1, Appl
22	15.2	69.1	7745	US-08-874-162-1	Sequence 1, Appl
23	15.2	69.1	1016	US-08-399-986B-3	Sequence 3, Appl
24	15.2	69.1	1016	US-08-493-754A-3	Sequence 3, Appl
25	15.2	69.1	2182	US-08-399-986B-1	Sequence 1, Appl
26	15.2	69.1	2182	US-08-493-754A-1	Sequence 1, Appl
27	15.2	69.1	2598	US-08-745-892-20	Sequence 20, Appl

C 28	15.2	69.1	5420	6	5256642-3	Patent No. 5256642
C 29	15.2	69.1	5420	6	5472939-3	Patent No. 5472939
C 30	15.2	69.1	5872	4	US-09-102-528-32	Sequence 32, Appl
C 31	15.2	69.1	6951	6	5256642-1	Patent No. 5256642
C 32	15.2	69.1	6951	6	5472939-1	Patent No. 5472939
C 33	14.8	67.3	337	4	US-08-991-789A-13	Sequence 13, Appl
C 34	14.8	67.3	1855	3	US-08-961-083-71	Sequence 71, Appl
C 35	14.8	67.3	3159	1	US-08-119-361-4	Sequence 4, Appl
C 36	14.8	67.3	3159	3	US-08-336-308A-3	Sequence 3, Appl
C 37	14.8	67.3	3159	3	US-08-822-324-3	Sequence 3, Appl
C 38	14.8	67.3	3159	4	US-09-490-931-3	Sequence 3, Appl
C 39	14.8	67.3	7266	3	US-08-336-308A-9	Sequence 9, Appl
C 40	14.8	67.3	7266	3	US-08-822-324-5	Sequence 5, Appl
C 41	14.8	67.3	7266	4	US-09-490-931-9	Sequence 9, Appl
C 42	14.8	67.3	8640	1	US-08-570-311-28	Sequence 28, Appl
C 43	14.8	67.3	9919	3	US-08-880-179-1	Sequence 1, Appl
C 44	14.6	66.4	293	3	US-08-866-340-13	Sequence 13, Appl
C 45	14.6	66.4	293	4	US-09-103-875-17	Sequence 17, Appl

ALIGNMENTS

```
RESULT 1
US-09-102-528-31/c
; Sequence 31, Application US/09102528
; Patent No. 6207883
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: DNA Sequences
; NUMBER OF SEQUENCES: 32
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/102,528
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/GB96/03191
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2582 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-102-528-31

Query Match      76.4%; Score 16.8; DB 4; Length 2582;
Best Local Similarity 90.0%; Pred. No. 14;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY      1 tcttgatcacacaagattga 20
Db      268 TGTGATCACACAGATCTGA 249

RESULT 2
US-08-014-326-1
; Sequence 1, Application US/08014326
; Patent No. 5498544
; GENERAL INFORMATION:
; APPLICANT: Gengenbach, Burle G.
; APPLICANT: Somers, David A.
; APPLICANT: Wyse, Donald L.
; APPLICANT: Gronwald, John W.
; APPLICANT: Egli, Margaret A.
; APPLICANT: Lutz, Shiela M.
; TITLE OF INVENTION: Method and An Acetyl CoA Carboxylase
```

```

; TITLE OF INVENTION: Alteration in Oil Content of Plants
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant & Gould
; STREET: 3100 No. 5498544west Center
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/014,326
; FILING DATE: 05-FEB-1993
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/917,462
; FILING DATE: 21-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/538,674
; FILING DATE: 18-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/269,584
; FILING DATE: 10-NOV-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Kowalczyk, Katherine M.
; REGISTRATION NUMBER: 36,848
; REFERENCE/DOCKET NUMBER: 600.258-US01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-332-5300
; TELEFAX: 612-332-9081
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2000 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: 2 kb fragment of lambda clone #15-14
; US-08-014-326-1

Query Match 73.6%; Score 16.2; DB 1; Length 2000;
Best Local Similarity 85.7%; Pred. No. 27;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 2 gtggatcacagaattgata 22
||||| ||||| ||| ||
Db 42 GTTGATGACAGAGTTGTTA 62

RESULT 3
US-08-417-089-1
; Sequence 1, Application US/08417089
; Patent No. 6069298
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: METHODS AND AN ACETYL COA CARBOXYLASE GENE
; TITLE OF INVENTION: FOR CONFERRING HERBICIDE TOLERANCE AND AN ALTERATION IN
; TITLE OF INVENTION: OIL CONTENT OF PLANTS
; NUMBER OF SEQUENCES: 11
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/417,089
; INFORMATION FOR SEQ ID NO: 1:
```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 2001 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-417-089-1

Query Match 73.6%; Score 16.2; DB 3; Length 2001;
Best Local Similarity 85.7%; Pred. No. 27;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 2 gtggatcacagaattgata 22
||||| ||||| ||| ||
Db 42 GTTGATGACAGAGTTGTTA 62

RESULT 4
US-08-695-651-1
; Sequence 1, Application US/08695651
; Patent No. 6146867
; GENERAL INFORMATION:
; APPLICANT: Gengenbach, B. G.
; APPLICANT: Somers, D. A.
; APPLICANT: Wyse, D. L.
; APPLICANT: Gronwald, J. W.
; APPLICANT: Egli, M. A.
; APPLICANT: Lutz, S. M.
; TITLE OF INVENTION: METHOD FOR PRODUCING PLANT ACETYL COA
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTED Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/695,651
; FILING DATE: 12-AUG-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/417089
; FILING DATE: 05-APR-1995
; APPLICATION NUMBER: 08/014326
; FILING DATE: 05-FEB-1993
; APPLICATION NUMBER: 07/917462
; FILING DATE: 21-JUL-1992
; APPLICATION NUMBER: 07/538674
; FILING DATE: 18-JUN-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Woessner, Warren D
; REGISTRATION NUMBER: 30,440
; REFERENCE/DOCKET NUMBER: 600.318US3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-373-6900
; TELEFAX: 612-339-3061
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2001 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; US-08-695-651-1
```

Query Match 73.6%; Score 16.2; DB 3; Length 2001;
Best Local Similarity 85.7%; Pred. No. 27;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 gttgagtcacaagatttgata 22
||||| ||||| ||| ||
DB 42 GTTGATGACACAGATTGTGA 62

RESULT 5
US-08-930-285-1
; Sequence 1, Application US/08930285
; Patent No. 6222099
; GENERAL INFORMATION:
; APPLICANT: Regents of the University of Minnesota, et al.
; TITLE OF INVENTION: TRANSGENIC PLANTS EXPRESSING ACETYL COA CARBOXYLASE GE
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P. A.
; STREET: P. O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/930,285
; FILING DATE: 13-APR-1998
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/04625
; FILING DATE: 04-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Embretson, Janet E.
; REGISTRATION NUMBER: 39,665
; REFERENCE/DOCKET NUMBER: 600.318054
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-339-0331
; TELEFAX: 612-339-3061
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2001 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; US-08-930-285-1

Query Match 73.6%; Score 16.2; DB 4; Length 2001;
Best Local Similarity 85.7%; Pred. No. 27;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 gttgagtcacaagatttgata 22
||||| ||||| ||||| ||
DB 42 GTTGATGACACAGATTGTGA 62

RESULT 6
US-08-695-421-1
; Sequence 1, Application US/08695421
; Patent No. 6268550
; GENERAL INFORMATION:

APPLICANT: Gengenbach, B. G.
; APPLICANT: Somers, D. A.
; APPLICANT: Wyse, D. L.
; APPLICANT: Gronwald, J. W.
; APPLICANT: Egli, M. A.
; APPLICANT: Lutz, S. M.
; TITLE OF INVENTION: METHODS AND AN ACETYL COA CARBOXYLASE GENE
; TITLE OF INVENTION: FOR CONFERRING HERBICIDE TOLERANCE AND AN ALTERATION IN OIL
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/695,421
; FILING DATE: 23-AUG-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/417089
; FILING DATE: 05-APR-1995
; APPLICATION NUMBER: 08/014326
; FILING DATE: 05-FEB-1993
; APPLICATION NUMBER: 07/917462
; FILING DATE: 21-JUL-1992
; APPLICATION NUMBER: 07/538674
; FILING DATE: 18-JUN-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Woessner, Warren D.
; REGISTRATION NUMBER: 30,440
; REFERENCE/DOCKET NUMBER: 600.318052
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-373-6900
; TELEFAX: 612-339-3061
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2001 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-695-421-1

Query Match 73.6%; Score 16.2; DB 4; Length 2001;
Best Local Similarity 85.7%; Pred. No. 27;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 gttgagtcacaagatttgata 22
||||| ||||| ||||| ||
DB 42 GTTGATGACACAGATTGTGA 62

RESULT 7
US-08-244-537-1
; Sequence 1, Application US/08244537
; Patent No. 5854420
; GENERAL INFORMATION:
; APPLICANT: ASHTON, ANTHONY R.
; APPLICANT: JENKINS, COLIN L.D.
; APPLICANT: WHITFIELD, PAUL R.
; TITLE OF INVENTION: MAIZE ACETYL COA CARBOXYLASE ENCODING
; TITLE OF INVENTION: DNA CLONES
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:

ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P.
STREET: 1100 New York Avenue, N.W.
CITY: Washington
STATE: D. C.
COUNTRY: U.S.A.
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/244,537
FILING DATE: 18-AUG-1994
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9125330.2
FILING DATE: 28-NOV-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB92/02205
FILING DATE: 27-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: KOKULIS, PAUL N.
REGISTRATION NUMBER: 16,773
REFERENCE/DOCKET NUMBER: 203094/SEE 36663/UST
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4345 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-244-537-1

Query Match 73.6%; Score 16.2; DB 2; Length 4345;
Best Local Similarity 85.7%; Pred. No. 30;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 gtggatcacagaattgata 22
||||| ||||| ||| |||
Db 902 GTGGATGACAAGACTGTCTTA 922

RESULT 8
US-08-417-089-5
Sequence 5, Application US/08417089
Patent No. 6069298
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: METHODS AND AN ACETYL COA CARBOXYLASE GENE
TITLE OF INVENTION: FOR CONFERRING HERBICIDE TOLERANCE AND AN ALTERATION IN
NUMBER OF SEQUENCES: 11
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/417,089
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 7470 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-417-089-5

Query Match 73.6%; Score 16.2; DB 3; Length 7470;
Best Local Similarity 85.7%; Pred. No. 32;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 gtggatcacagaattgata 22
||||| ||||| ||| |||
Db 3972 GTTGATGACAAGACTGTCTTA 3992

RESULT 9
US-08-695-651-5
Sequence 5, Application US/08695651
Patent No. 6146867
GENERAL INFORMATION:
APPLICANT: Gengenbach, B. G.
APPLICANT: Somers, D. A.
APPLICANT: Wyse, D. L.
APPLICANT: Gronwald, J. W.
APPLICANT: Egli, M. A.
APPLICANT: Lutz, S. M.
TITLE OF INVENTION: METHOD FOR PRODUCING PLANT ACETYL COA
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
STREET: P.O. Box 2938
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55402
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/695,651
FILING DATE: 12-AUG-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/417089
FILING DATE: 05-APR-1995
APPLICATION NUMBER: 08/014326
FILING DATE: 05-FEB-1993
APPLICATION NUMBER: 07/917462
FILING DATE: 21-JUL-1992
APPLICATION NUMBER: 07/538674
FILING DATE: 18-JUN-1990
ATTORNEY/AGENT INFORMATION:
NAME: Woessner, Warren D
REGISTRATION NUMBER: 30,440
REFERENCE/DOCKET NUMBER: 600.318US3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-373-6900
TELEFAX: 612-339-3061
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 7470 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-695-651-5

Query Match 73.6%; Score 16.2; DB 3; Length 7470;
Best Local Similarity 85.7%; Pred. No. 32;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 gtggatcacagaattgata 22
||||| ||||| ||| |||

Db 3972 GTTGATGACAGAGTTGTTA 3992

RESULT 10

US-08-930-285-5

Sequence 5, Application US/08930285

Patent No. 6222099

GENERAL INFORMATION:

APPLICANT: Regents of the University of Minnesota, et al.

TITLE OF INVENTION: TRANSGENIC PLANTS EXPRESSING ACETYL COA CARBOXYLASE GE

NUMBER OF SEQUENCES: 26

CORRESPONDENCE ADDRESS:

ADDRESSEE: Schweigman, Lundberg, Moessner & Kluth, P. A.

STREET: P. O. Box 2938

CITY: Minneapolis

STATE: MN

COUNTRY: USA

ZIP: 55402

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/930,285

FILING DATE: 13-APR-1998

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US96/04625

FILING DATE: 04-APR-1996

ATTORNEY/AGENT INFORMATION:

NAME: Embretson, Janet E.

REGISTRATION NUMBER: 39,665

REFERENCE/DOCKET NUMBER: 600.318054

TELECOMMUNICATION INFORMATION:

TELEPHONE: 612-339-0331

TELEFAX: 612-339-3061

TELEX:

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 7470 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE:

ORIGINAL SOURCE:

US-08-930-285-5

Query Match 73.6%; Score 16.2; DB 4; Length 7470;

Best Local Similarity 85.7%; Pred. No. 32;

Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 gttgatcacagaattgata 22

Db 3972 GTTGATGACAGAGTTGTTA 3992

RESULT 11

US-08-695-421-5

Sequence 5, Application US/08695421

Patent No. 6258550

GENERAL INFORMATION:

APPLICANT: Gengenbach, B. G.

APPLICANT: Somers, D. A.

APPLICANT: Wyse, D. L.

APPLICANT: Gronwald, J. W.

APPLICANT: Egli, M. A.

APPLICANT: Lutz, S. M.

TITLE OF INVENTION: METHODS AND AN ACETYL COA CARBOXYLASE GENE

TITLE OF INVENTION: FOR CONFERRING HERBICIDE TOLERANCE AND AN ALTERATION IN OIL

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: Schweigman, Lundberg, Moessner & Kluth, P. A.

STREET: P. O. Box 2938

CITY: Minneapolis

STATE: MN

COUNTRY: USA

ZIP: 55402

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/695,421

FILING DATE: 23-AUG-1996

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/417089

FILING DATE: 05-APR-1995

APPLICATION NUMBER: 08/014326

FILING DATE: 05-FEB-1993

APPLICATION NUMBER: 07/917462

FILING DATE: 21-JUL-1992

APPLICATION NUMBER: 07/538674

FILING DATE: 18-JUN-1990

ATTORNEY/AGENT INFORMATION:

NAME: Moessner, Warren D

REGISTRATION NUMBER: 30,440

REFERENCE/DOCKET NUMBER: 600.318052

TELECOMMUNICATION INFORMATION:

TELEPHONE: 612-373-6900

TELEFAX: 612-339-3061

TELEX:

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 7470 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

US-08-695-421-5

Query Match 73.6%; Score 16.2; DB 4; Length 7470;

Best Local Similarity 85.7%; Pred. No. 32;

Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 gttgatcacagaattgata 22

Db 3972 GTTGATGACAGAGTTGTTA 3992

RESULT 12

US-08-602-145-14/C

Sequence 14, Application US/08602145

Patent No. 6025336

GENERAL INFORMATION:

APPLICANT: Goltzy, Kristin L.

APPLICANT: Greenberger, Joel S.

TITLE OF INVENTION: DETERMINING EXPOSURE TO IONIZING RADIATION

TITLE OF INVENTION: AGENT WITH PERSISTENT BIOLOGICAL MARKERS

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: Foley & Lardner

STREET: 3000 K Street, N.W., Suite 500

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20007-5109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/602,145
FILING DATE: 15-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: Bent, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 076333/0112
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 296 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-602-145-14

Query Match 70.9%; Score 15.6; DB 3; Length 296;
Best Local Similarity 81.8%; Pred. No. 42;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ttttgatcacagaatttgata 22
|||||
DB 288 TTTTCATCATATACATTGTGATA 267

RESULT 13
US-08-928-383B-3/c
Sequence 3, Application US/8928383B
Patent No. 6210921
GENERAL INFORMATION:
APPLICANT: Robert W. Finberg, Jeffrey M. Bergelson,
APPLICANT: and Marshall S. Horwitz
TITLE OF INVENTION: CAR, A No. 6210921el Cocksacklevirus and Adenovirus
TITLE OF INVENTION: Receptor
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/928,383B
FILING DATE: 12-SEP-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/026,100
FILING DATE: 13-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: DFN-020
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1095 base pairs
TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1095
US-08-928-383B-3

Query Match 70.9%; Score 15.6; DB 4; Length 1095;
Best Local Similarity 81.8%; Pred. No. 49;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ttttgatcacagaatttgata 22
|||||
DB 312 TGATGATCACACGATTGTGAGA 291

RESULT 14
US-08-684-862-8/c
Sequence 8, Application US/08684862
Patent No. 5759541
GENERAL INFORMATION:
APPLICANT: Bach, Alfred
APPLICANT: Hilleu, Heinz
APPLICANT: Bialojan, Siegfried
TITLE OF INVENTION: No. 5759541el Proteins, the Preparation and Use
TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Keil & Weinkauff
STREET: 1101 Connecticut Avenue
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
COMPUTER: IBM AT-compatible, 80286 processor
OPERATING SYSTEM: MS-DOS version 5.0
SOFTWARE: Wordperfect version 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/684,862
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/361,705
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,040
FILING DATE: 30-DEC-1992
APPLICATION NUMBER: PCT/EP91/01361
FILING DATE: 19-JUL-1991
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1096 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: Agkistrodon rhodostoma
FEATURE:
LOCATION: 144 to 841
OTHER INFORMATION: the coding region shown in (2)(1x)(B)
OTHER INFORMATION: codes for the protein of SEQ ID NO: 3
US-08-684-862-8

Query Match 70.9%; Score 15.6; DB 1; Length 1096;
Best Local Similarity 81.8%; Pred. No. 49;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 tgttgatcacagaatttgata 22
||| | ||||| |||| |
DB 112 TGTAGATCACAAGTTTCTA 91

RESULT 15

US-08-928-383B-1/c
; Sequence 1, Application US/08928383B
; Patent No. 6210921
; GENERAL INFORMATION:
; APPLICANT: Robert W. Finberg, Jeffrey M. Bergelson,
; APPLICANT: and Marshall S. Horwitz
; TITLE OF INVENTION: CAR, A No. 6210921el Cocksacklevirus and Adenovirus
; TITLE OF INVENTION: Receptor
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/928,383B
; FILING DATE: 12-SEP-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/026,100
; FILING DATE: 13-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragoras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: DFN-020
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1584 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 60..1157
; US-08-928-383B-1

Query Match 70.9%; Score 15.6; DB 4; Length 1584;
Best Local Similarity 81.8%; Pred. No. 51;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
OY 1 tgttgatcacagaatttgata 22
||| | ||||| |||| |
DB 371 TGATGCATCACCGAGATTGAGA 350

Search completed: December 15, 2001, 03:22:46
Job time: 7561 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 02:33:32 ; Search time 4441.54 Seconds
(without alignments)
53.226 Million cell updates/sec

Title: US-09-380-826a-2

Sequence: 1 ttttgatcacacagattgata 22

Scoring table: IDENTITY_NUC
Gapop 10.0 ; Gapext 1.0

Searched: 11351937 seqs, 537289281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_hic:*
10: gb_est1:*
11: gb_est2:*
12: gb_hic:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rpd:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	18.8	85.5	361	10	AV532040 AV532040
C 2	18.8	85.5	394	11	226559 ATG51651 Ve
C 3	18.8	85.5	490	10	AB038725 AB038725
C 4	18.8	85.5	527	10	AV520789 AV520789
C 5	18.8	85.5	530	10	AV520789 AV520789
C 6	18.4	83.6	608	11	BE977618 bs6b08.y
C 7	18.4	83.6	934	13	CNS02BWA
C 8	17.8	80.9	600	11	BF006444 EST434942
C 9	17.8	80.9	633	11	BF006254 EST434752
C 10	17.4	79.1	286	10	BR288727 BR288727
C 11	17.4	79.1	407	10	BE581720 Kq51c07.y
C 12	17.4	79.1	529	13	A0497779 HS_5066_B

C 13	17.4	79.1	646	10	AW695022 NF082E04S
C 14	17.4	79.1	671	11	BC448853 NF003C10I
C 15	17.4	79.1	762	10	BE283043 601101323
C 16	17.4	79.1	837	10	BE642809 Crl2_7_B0
C 17	17.4	79.1	1101	13	CNS00DB2
C 18	17.2	78.2	151	10	BE068225 MR4-B7036
C 19	17.2	78.2	425	10	A1049911 an30h03.x
C 20	17.2	78.2	431	13	AO595814 HS_2132_B
C 21	17.2	78.2	432	13	AO221715 HS_2010_A
C 22	17.2	78.2	437	11	W43212 22591_Lambda
C 23	17.2	78.2	439	13	AO796989 nbx0071H
C 24	17.2	78.2	500	10	A1733664 an30h03.x
C 25	17.2	78.2	524	13	B62582 T22F18TR_TA
C 26	17.2	78.2	531	13	AO223477 HS_2003_B
C 27	17.2	78.2	536	10	AO034253 EST277824
C 28	17.2	78.2	546	10	A1779714 EST260593
C 29	17.2	78.2	579	13	B67515 T22M10TR_TA
C 30	17.2	78.2	602	10	A1777095 EST258060
C 31	17.2	78.2	614	10	BE187570 EST336131
C 32	17.2	78.2	619	10	AW963686 EST375759
C 33	17.2	78.2	630	13	BF346432 602020273
C 34	17.2	78.2	680	13	B57784 C1T-HSP-201
C 35	17.2	78.2	698	13	A2193715 Sp_1023_B
C 36	17.2	78.2	773	13	CNS07DUS
C 37	17.2	78.2	888	13	CNS079F4
C 38	17.2	78.2	904	13	CNS029DN
C 39	17.2	78.2	949	13	CNS04HNP
C 40	17	77.3	373	11	C70416
C 41	17	77.3	648	10	A0200312
C 42	16.8	76.4	238	10	AV313332 AV313332
C 43	16.8	76.4	260	10	BE118999 UI-R-CA0-
C 44	16.8	76.4	296	11	T20362 6c01908-C7
C 45	16.8	76.4	308	11	BI398487 952002A03

ALIGNMENTS

RESULT 1
LOCUS AV532040 361 bp mRNA EST 01-SEP-2000
DEFINITION AV532040 Arabidopsis thaliana flower buds Columbia Arabidopsis
thaliana cDNA clone FB034f09F 3', mRNA sequence.
ACCESSION AV532040
VERSION AV532040.1 GI:8692323
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 361)
Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries
JOURNAL DNA Res. 7, 175-180 (2000)
MEDLINE 20363093
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizukazusa.or.jp URL: http://www.kazusa.or.jp/en/plant/.
FEATURES
source
1..361
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="FB034f09F"
/lone_jib="Arabidopsis thaliana flower buds Columbia"
/tissue_type="flower buds"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT	113 a	73 c	61 g	114 t
ORIGIN				
Query Match		85.5%;	Score 18.8;	DB 10; Length 361;
Best Local Similarity		90.9%;	Pred. No. 99;	
Matches	20;	Conservative	0;	Mismatches 2; Indels 0; Gaps 0;
Oy	1	tgttgatcacagaatttgata	22	
Db	276	tgttgcttcacagaatttgaca	255	
RESULT 2				
LOCUS	226559	394 bp	mRNA	EST 31-MAY-1995
DEFINITION	ATP5I651 Versailles-VC Arabidopsis thaliana cDNA clone VCVDH08 3'			
ACCESSION	226559			
VERSION	226559.1	GI:404227		
KEYWORDS	EST.			
SOURCE	thale cress.			
ORGANISM	Arabidopsis thaliana			
REFERENCE	Arabidopsis thaliana transcribed genome: the GDR cDNA program			
AUTHORS	Unpublished (1996)			
TITLE	Contact: Desprez T., Amselem J., Chiapello H., Rouze P., Caboche			
JOURNAL	M., Hofte H.			
COMMENT	INRA Versailles			
	Laboratoire de Biologie Cellulaire			
	Route de Saint-Cyr, 78026 Versailles Cedex, France			
	Email: thierry@versailles.inra.fr.			
FEATURES	Location/Qualifiers			
source	1..394			
	/organism="Arabidopsis thaliana"			
	/strain="ecotype Columbia"			
	/db_xref="taxon:3702"			
	/clone="VCVDH08"			
	/clone_lib="Versailles-VC"			
	/tissue_type="whole seedlings"			
	/dev_stage="in vitro-grown etiolated seedlings, 5 days old"			
	/note="Vector: pBluescript"			
	117 a 80 c 63 g 133 t 1 others			
BASE COUNT	117 a	80 c	63 g	133 t
ORIGIN				
Query Match		85.5%;	Score 18.8;	DB 11; Length 394;
Best Local Similarity		90.9%;	Pred. No. 1e+02;	
Matches	20;	Conservative	0;	Mismatches 2; Indels 0; Gaps 0;
Oy	1	tgttgatcacagaatttgata	22	
Db	321	tgcttgcttcacagaatttgaca	300	
RESULT 3				
LOCUS	AB038725	490 bp	mRNA	EST 14-NOV-2000
DEFINITION	AB038725 Arabidopsis thaliana Above-ground organ from two to six-week old plants Columbia Arabidopsis thaliana cDNA clone			
ACCESSION	AB038725			
VERSION	AB038725.1	GI:7212552		
KEYWORDS	EST.			
SOURCE	thale cress.			
ORGANISM	Arabidopsis thaliana			
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;			
	ROSIDAE; eurosids II; Brassicales; Brassicaceae; Arabidopsids.			

REFERENCE	1 (bases 1 to 490)									
AUTHORS	Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.									
TITLE	A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries									
JOURNAL	DNA Res. 7, 175-180 (2000)									
MEDLINE	20363093									
COMMENT	Contact: Erika Asamizu The First Laboratory for Plant Gene Research Kazusa DNA Research Institute Yana 1532-3, Kisarazu, Chiba 292-0812, Japan Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/ Location/Qualifiers									
FEATURES	source									
source	1..490									
	/organism="Arabidopsis thaliana"									
	/strain="Columbia"									
	/db_xref="taxon:3702"									
	/clone_1ib="Ap230603_1"									
	/clone_1ib="Arabidopsis thaliana Above-ground organ from two to six-week old plants Columbia"									
	/tissue.type="aboveground organ"									
	/dev_stage="two to six-week old plants"									
	/note="Vector: pBluescriptII SK-, Site_1: EcoRI; Site_2: XhoI"									
BASE COUNT	147 a 99 c 87 g 157 t									
ORIGIN										
Query Match	85.5%; Score 18.8; DB 10; Length 490;									
Best Local Similarity	90.9%; Pred. No.1e+02; 2; Indels 0; Gaps 0;									
Matches	20; Conservative 0; Mismatches 20									
Qy	1 tgttgatcacaagaatttgata 22									
Db	303 TGTGTGTTCCACAGATTGACA 282									
RESULT 4										
LOCUS	AV520789 527 bp mRNA EST 01-SEP-2000									
DEFINITION	AV520789 Arabidopsis thaliana aboveground organs two to six-week old Arabidopsis thaliana cDNA clone AP232906F 3', mRNA sequence.									
ACCESSION	AV520789									
VERSION	AV520789.1 GI:8680316									
KEYWORDS	EST.									
SOURCE	thale cress.									
ORGANISM	Arabidopsis thaliana									
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.									
REFERENCE	1 (bases 1 to 527)									
AUTHORS	Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.									
TITLE	A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries									
JOURNAL	DNA Res. 7, 175-180 (2000)									
MEDLINE	20363093									
COMMENT	Contact: Erika Asamizu The First Laboratory for Plant Gene Research Kazusa DNA Research Institute Yana 1532-3, Kisarazu, Chiba 292-0812, Japan Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/ Location/Qualifiers									
FEATURES	1..527									
source	/organism="Arabidopsis thaliana"									
	/strain="Columbia"									
	/db_xref="taxon:3702"									
	/clone="AP232906F"									
	/clone_1ib="Arabidopsis thaliana aboveground organs two to six-week old"									
	/tissue.type="aboveground organs"									
	/dev_stage="two to six-week old"									
	/note="Vector: pBluescriptII SK-, Site_1: EcoRI; Site_2:									

BASE COUNT 157 a 111 c 101 g 158 t
 ORIGIN

Query Match 85.5%; Score 18.8; DB 10; Length 527;
 Best Local Similarity 90.9%; Pred. No. 1.1e+02;
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ttttgatcacaaagatttgata 22
 ||||| ||||| ||||| ||||| |||||
 Db 247 TGTGTGTTCAACAAGATTGACA 226

RESULT 5
 A1995637/c 530 bp mRNA EST 08-SEP-1999
 LOCUS 701676626 A. thaliana, Columbia Col-0, inflorescence-1 Arabidopsis
 DEFINITION thaliana cDNA clone 701676626, mRNA sequence.
 ACCESSION A1995637
 VERSION A1995637.1 GI:5842542
 KEYWORDS EST
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 530)
 Chen, J., Momiyama, M., Chan, E., Mooney, M., Carroon, B., Gilliland, D.,
 Wang, X., Hillman, J., Guegler, K., Kim, C., Doyle, M., Brzoska, P.,
 Gorgone, G., Burns, D., Griffin, J., Mouanoutoua, M., Nguyen, D., Tan, R.,
 Rose, M., Warren, B., Ton, B., Kastury, K., Borillo, C., Carpilo, T.,
 Policky, J., Suzuki, G., Argentine, C., Shah, S., Nohrrega, A., Murry, L.,
 Turner, C., Kirkorian, S., Elder, L. and Hanson, D.
 Arabidopsis thaliana Gene Expression Microarray
 Unpublished (1999)
 CONTACT: David Smoller, Ph.D.
 Genome Systems, Inc., a wholly owned subsidiary of Incyte
 Pharmaceuticals, Inc.
 4633 World Parkway Circle, St. Louis, MO 63134, USA
 Tel: 877-577-2733
 Fax: 314-427-3324
 Email: service@genomesystems.com.
 Location/Qualifiers

FEATURES
 source 1..530
 /organism="Arabidopsis thaliana"
 /cultivar="Columbia Col-0"
 /db_xref="taxon:3702"
 /clone="701676626"
 /clone_lib="A. thaliana, Columbia Col-0, inflorescence-1"
 /tissue_type="inflorescence"
 /dev_stage="4 - 7 weeks"
 /note="Vector: pSPORT; Site_1: NotI; Site_2: SalI; cDNA
 library was derived from untreated inflorescence tissue
 from Arabidopsis thaliana, Columbia Col-0, at 4 - 7
 weeks. Plants were grown in 1:1:1 peat
 moss/vermiculite/perlite soil at 22 deg. C +/- 3 deg. C
 under constant light, and watered with fertilizer. cDNA
 synthesis was initiated using a NotI-oligo(dT) primer.
 Double-stranded cDNA was blunt-ended, ligated to SalI adaptors
 digested with NotI, size-selected, and cloned into the
 NotI and SalI sites of the pSPORT vector."

BASE COUNT 163 a 108 c 100 g 158 t 1 others
 ORIGIN

Query Match 85.5%; Score 18.8; DB 10; Length 530;
 Best Local Similarity 90.9%; Pred. No. 1.1e+02;
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ttttgatcacaaagatttgata 22
 ||||| ||||| ||||| ||||| |||||
 Db 254 TGTGTGTTCAACAAGATTGACA 233

RESULT 6

BE977618/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BE977618 608 bp mRNA EST 04-OCT-2000
 bs66h08.y1 Drosophila melanogaster adult testis library Drosophila
 melanogaster cDNA clone bs66h08 5', mRNA sequence.
 BE977618
 BE977618.1 GI:10608272
 EST.
 Fruit fly.
 Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 608)
 Andrews, J., Bouffard, G. and Oliver, B.
 Drosophila melanogaster testis expressed sequence tags
 Unpublished (1999)
 CONTACT: Brian Oliver
 Laboratory of Cellular and Developmental Biology
 NIDDK, National Institutes of Health
 6 Center Drive MSC 2715, Bldg 6, Rm B1-13, Bethesda, MD 20892 USA
 Fax: (301) 496 5239
 Email: oliverhelix.nih.gov,
 http://www.nidk.nih.gov/intram/people/boliver.htm
 Tissue isolation and library construction performed at the National
 Institute of Diabetes and Digestive and Kidney Diseases, NIH (see
 http://www.nidk.nih.gov/intram/people/boliver.htm). DNA sequencing
 and analyses performed by National Institutes of Health Intramural
 Sequencing Center (NISC; see http://www.nisc.nih.gov).
 Plate: 66 row: h column: 08
 Seq primer: M13rpl reverse primer (AB1).
 Location/Qualifiers
 1..608
 /organism="Drosophila melanogaster"
 /strain="y1 w[67c1]/Y"
 /db_xref="taxon:7227"
 /clone="bs66h08"
 /clone_lib="Drosophila melanogaster adult testis library"
 /sex="male"
 /dev_stage="1-5 day adult"
 /lab_host="SOLR (Stratagene)"
 /note="Organ: testis; Vector: pBluescript SK (Stratagene);
 Site_1: EcoR I; Site_2: Xho I; Testes dissected from 1-5
 day adult y1 w[67c1]/Y males raised at 25°C. RNA
 isolated using Trizol (Life Technologies) and a single
 round of Poly(A)+ selection using Oligotex (Qiagen). cDNA
 library constructed using Stratagene Zap-cDNA synthesis
 kit. Oligo dT-primed, size fractionated -1.6 kb, and
 directionally cloned at EcoRI and XhoI in uni-zap XR.
 Following a single round of amplification pBluescript SK
 phagemids were mass excised. A distribution channel for
 clones is being sought, but not currently available.
 Requests for clones cannot be honored."

BASE COUNT 173 a 146 c 161 g 128 t
 ORIGIN

Query Match 83.6%; Score 18.4; DB 11; Length 608;
 Best Local Similarity 95.0%; Pred. No. 1.7e+02;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 tttgatcacaaagatttgata 22
 ||||| ||||| ||||| ||||| |||||
 Db 74 TTGATGCCAAGATTGATA 55

RESULT 7
 CANS02BMA/c 934 bp DNA GSS 12-MAY-2000
 LOCUS CANS02BMA/c 934 bp DNA GSS 12-MAY-2000
 DEFINITION Tetradon nigroviridis genome survey sequence PUC-ori end of clone
 254C22 of library G from Tetradon nigroviridis, genomic survey

```

sequence.
ACCESSION      AL190387.1  GI:7828491
VERSION        GSS; genome survey sequence.
KEYWORDS       Tetraodon nigroviridis.
SOURCE         Tetraodon nigroviridis
ORGANISM       Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Acanthopterygii; Neopterygii; Teleostei; Euteleostei;
               Tetraodontidae; Tetraodon.
REFERENCE      1 (bases 1 to 934)
AUTHORS        Roest-Crolius,H., Jaillon,O., Dasilva,C., Fitzames,C., Fisher,C.,
               Bouneau,L., Billault,A., Queller,F., Saurin,W., Bernot,A. and
               Weissenbach,J.
TITLE          Characterization and repeat analysis of the compact genome of the
               freshwater pufferfish Tetraodon nigroviridis
JOURNAL        Unpublished
REFERENCE      2 (bases 1 to 934)
AUTHORS        Roest-Crolius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
               Bernot,A., Fitzames,C., Mincker,P., Brottier,P., Queller,F.,
               Saurin,W. and Weissenbach,J.
TITLE          Human gene number estimate provided by genome wide analysis using
               Tetraodon nigroviridis DNA sequence
JOURNAL        Unpublished
REFERENCE      3 (bases 1 to 934)
AUTHORS        Direct Submission
COMMENT        Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases
               This sequence is a single read and was generated as part of a large
               scale clone-end sequencing project of the Tetraodon nigroviridis
               genome. For more information, please take a look at
               http://www.genoscope.cns.fr/Tetraodon.
               Location/Qualifiers
FEATURES       1..934
               /organism="Tetraodon nigroviridis"
               /db_xref="taxon:99883"
               /clone="254C22"
               /clone_1db="G"
               /note="Genoscope sequence ID : C0AG254B811SP1-end :
               PUC-ori"
BASE COUNT     282 a 175 c 207 g 259 t 11 others
ORIGIN
Query Match      83.6%; Score 18.4; DB 13; Length 934;
Best Local Similarity 95.0%; Pred. No. 1.8e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 3 ttgatcacaagaattgata 22
||||| |||||||||
Db 840 TTGGAACACAGATTGTGATA 821

RESULT 8
BF006444/c 600 bp mRNA EST 06-OCT-2000
LOCUS      EST434942 DSLC Medicago truncatula cDNA clone pDSLc-41J15, mRNA
DEFINITION sequence.
ACCESSION  BF006444
VERSION    BF006444.1 GI:10706719
KEYWORDS   EST.
SOURCE     barrel medic.
ORGANISM   Medicago truncatula
               Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
               Rosidae; eucosids I; Fabales; Fabaceae; Papilionoidae; Trifoliaceae;
               Medicago.
               1 (bases 1 to 600)
               Fedorova,M., Pierson,B.L., Samac,D.A., Gantt,J.S., Vance,C.P.,
               Gonzales,M.B. and Ellis,L.
               ESTs from Medicago truncatula leaves and cotyledons
               Unpublished (2000)
               Contact: Deborah A. Samac
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Contact: Deborah A. Samac

```

```

Department of Plant Pathology
University of Minnesota
495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA
Tel: 612 625 1243
Fax: 651 649 5058
Email: debbys@puccini.crl.umn.edu
University of Minnesota name: M275482e TIGR sequence name:
MTLBA567K More information is available at:
http://chrysis.tamu.edu/medicago
Seq primer: Skmod (CTA gaa gta gat CC).
Location/Qualifiers
FEATURES       1..600
               /organism="Medicago truncatula"
               /cultivar="genotype A17"
               /db_xref="taxon:3880"
               /clone="pDSLc-41J15"
               /clone_1db="DSLc"
               /tissue_type="leaves and cotyledons"
               /dev_stage="mixture of cotyledons from five days old
               plants and leaves obtained from two weeks old plants"
               /lab_host="E. coli strain SOLR"
               /note="Vector: pluscript SK +/-; Site 1: EcoRI; Site 2:
               XhoI; cDNA was prepared from polyA+ enriched RNA from the
               mixture of cotyledons of five days old plants and leaves
               of two weeks old plants. The cDNA was directionally
               ligated into the Uni-ZAP XR vector from Stratagene and
               packaged using Gigapack III Gold packaging extracts.
               Plasmids containing cDNA inserts were excised from the
               recombinant lambda-ZAP phage using Ex-Assist helper phage
               and propagated in SOLR cells."
BASE COUNT     127 a 120 c 132 g 221 t
ORIGIN
Query Match      80.9%; Score 17.8; DB 11; Length 600;
Best Local Similarity 90.5%; Pred. No. 3.2e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 ttgtgatcacaagaattgat 21
||||| |||||||||
Db 90 TGTGATCACAAGATTGTGAT 70

RESULT 9
BF006254/c 633 bp mRNA EST 06-OCT-2000
LOCUS      EST434752 DSLC Medicago truncatula cDNA clone pDSLc-40G21, mRNA
DEFINITION sequence.
ACCESSION  BF006254
VERSION    BF006254.1 GI:10706529
KEYWORDS   EST.
SOURCE     barrel medic.
ORGANISM   Medicago truncatula
               Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
               Rosidae; eucosids I; Fabales; Fabaceae; Papilionoidae; Trifoliaceae;
               Medicago.
               1 (bases 1 to 633)
               Fedorova,M., Pierson,B.L., Samac,D.A., Gantt,J.S., Vance,C.P.,
               Gonzales,M.B. and Ellis,L.
               ESTs from Medicago truncatula leaves and cotyledons
               Unpublished (2000)
               Contact: Deborah A. Samac
               Department of Plant Pathology
               University of Minnesota
               495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA
               Tel: 612 625 1243
               Fax: 651 649 5058
               Email: debbys@puccini.crl.umn.edu
               University of Minnesota name: M275292e TIGR sequence name:
               MTLAU477K More information is available at:
               http://chrysis.tamu.edu/medicago
               Seq primer: Skmod (CTA gaa gta gat CC).
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

```

FEATURES
source
Location/Qualifiers
1. 633
/organism="Medicago truncatula"
/cultivar="genotype A17"
/db_xref="taxon:3880"
/clone="PDSLC-40621"
/clone_1ib="DSLIC"
/tissue_type="Leaves and cotyledons"
/dev_stage="mixture of cotyledons from five days old plants and leaves obtained from two weeks old plants"
/lab_host="E. coli strain SOLR"
/note="Vector: pBluescript SK +/-. Site_1: EcoRI, site_2: XhoI. cDNA was prepared from polyA+ enriched RNA from the mixture of cotyledons of five days old plants and leaves of two weeks old plants. The cDNA was directionally ligated into the Uni-ZAP XR vector from Stratagene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-ZAP phage using Ex-Assist helper phage and propagated in SOLR cells."

BASE COUNT
148 a 132 c 124 g 229 t

ORIGIN

Query Match
Best local Similarity 90.5%; Score 17.8; DB 11; Length 633;
Pred. No. 3.2e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 tcttgatcacagaattgat 21
||||| ||||||| |||

Db 188 TGTGATCATCACAGATTAGAT 168

RESULT 10
BB288727/c
LOCUS.
BB288727 286 bp mRNA EST 09-JUL-2000
DEFINITION BB288727 RIKEN full-length enriched, 2 cells egg mus musculus cDNA
clone B020032C12 3', mRNA sequence.
ACCESSION BB288727
VERSION BB288727.1 GI:8969176
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 286)
Komuro,H., Aizawa,K., Akahira,S., Akiyama,J., Arikawa,T., Carninci
,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N.,
Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M.,
Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N.,
Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M.,
Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y.,
Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata
,Y., Shigemoto,Y., Shingawa,A., Shiraki,T., Sogabe,Y., Sugahara,Y.,
Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tomimaga,N., Toya
,T., Tsunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yamana,I.,
Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino
,M., Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Komuro,H., et al.)
Unpublished (2000)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suenho-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-resgsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki
N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Thermolabile and thermolabile enzymes by
trehalose and its application for the synthesis of full length

CDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh,M., Kitunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,
Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki
,Y. and Hayashizaki,Y.
Automated filtration-based high-throughput plasmid preparation
system. Genome Res. 9 (5), 463-470 (1999)
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
19-44 (1999)
Please visit our web site (<http://genome.riken.go.jp>) for
further details.

FEATURES
source
Location/Qualifiers
1. 286
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="B020032C12"
/clone_1ib="RIKEN full-length enriched, 2 cells egg"
/tissue_type="egg"
/dev_stage="2 cells"
/lab_host="DH10B"
/note="Site_1: SalI; Site_2: BamHI. cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGACATCCAGACCTCTTTTCTTTTCTTTTCTTTN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence 15'
GAGAGAGACATCCAGCTTAATTAATTAATTCACCCGCCGCC 3'. cDNA
was cleaved with XhoI and BamHI. Vector: a modified
pBluescript KS(+) after bulk excision from Lambda FIC I."

BASE COUNT
117 a 66 c 20 g 83 t

ORIGIN

Query Match
Best local Similarity 94.7%; Score 17.4; DB 10; Length 286;
Pred. No. 4.3e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 gttgacacagaattga 20
||||| ||||||| |||

Db 169 GTTGATCATCACATTGGA 151

RESULT 11
BE581720/c
LOCUS.
BE581720 407 bp mRNA EST 09-MAY-2001
DEFINITION BE581720 K951C07.Y1 TBN95TH-SSR Strongyloides stercoralis cDNA 5' similar to
WP:119B10.2 CE16413 ;, mRNA sequence.
ACCESSION BE581720
VERSION BE581720.1 GI:9832662
KEYWORDS EST.
SOURCE Strongyloides stercoralis.
ORGANISM Strongyloides stercoralis.
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitda;
Panagrolaimidae; Strongyloididae; Strongyloides.
1 (bases 1 to 407)
McCartner,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J., Wylie,T.,
Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B., Bowers,Y.,
Gibbons,M., Ritter,E., Bennett,J., Franklin,C., Tsagarelisvili,R.,
Konko,I., Kennedy,S., Maguire,L., Beck,C., Underwood,K., Steptoe
,M., Allen,M., Person,B., Swaller,T., Harvey,N., Schurk,R., Kohn,S.,
Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and
Wilson,R.
The Washington Univ. Nematode EST Project, 1999
Unpublished (1999)
Contact: McCartner JP
The Washington Univ. Nematode EST Project, 1999
Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810

Email: estevenson.wustl.edu
 The library was constructed by Dr. Thomas Nutman and colleagues of
 NIAID, NIH (nutman@nih.gov). DNA Sequencing by: Washington
 University Genome Sequencing Center St. Louis.
 High quality sequence stop: 384.
 Location/Qualifiers

FEATURES

SOURCE

1. 407
 /organism="Strongyloides stercoralis"
 /strain="Rhabditiform larvae obtained from gerbils"
 /db_xref="taxon:6248"
 /clone_lib="JBN95TM-SSR"
 /lab_host="XL-1 Blue MRF" (Stratagene); Site: 1:
 /note="Vector: Lambda Uni-ZAP XR (Stratagene); Site: 1:
 EcoRI; Site: 2: XhoI; mRNA was purified from 2 x 10E3
 rhabditiform larvae which had been isolated from gerbils
 experimentally infected with larvae originally isolated
 from experimentally infected dogs. cDNA was constructed
 and, using adaptors, was cloned unidirectionally into the
 vector from the EcoRI site to the XhoI site. The library
 has an unamplified titer of 1 x 10E5 pfu/ml and an
 amplified, undiluted titer of 9 x 10E11 pfu/ml. The
 average insert size of the unamplified library is 675 bp
 (range, 100-1700)."
 152 a 63 c 74 g 118 t

BASE COUNT

ORIGIN

Query Match 79.1%; Score 17.4; DB 10: Length 407;
 Best Local Similarity 94.7%; Pred. No. 4.6e+02;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 ttggtcacagaatttgata 22
 1 |||||

Db 334 TTGATCACAAGATTGGATA 316

RESULT 12

LOCUS A0497779 529 bp DNA GSS 28-APR-1999
 DEFINITION HS_5066_B1.E05.T7A.RPCI-11 Human Male BAC Library Homo sapiens
 genomic clone Plate=642 COL=9 ROW=J, DNA sequence.

ACCESSION A0497779.1 GI:4697902
 VERSION A0497779.1
 KEYWORDS GSS.

SOURCE

ORGANISM

human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 529)
 Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
 Kellar,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
 Hood,L.

Sequence-tagged connectors: A sequence approach to mapping and
 scanning the human genome

Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)

TITLE

JOURNAL

MEDLINE

COMMENT

High Throughput Sequencing Center
 Contact: Mahairas GG, Wallace JC, Hood L
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu
 Clones are derived from the human BAC library RPCI-11. For BAC
 library availability, please contact Pieter de Jong
 (pieter@dejong.med.buflalo.edu). Clones may be purchased from
 BACPAC Resources (<http://bacpac.med.buflalo.edu/ordering/bac.htm>)
 or from Research Genetics (info@resgen.com). BAC end Web Server:
<http://www.htsc.washington.edu>
 Plate: 642 row: J column: 9

Seq primer: T7
 Class: BAC ends
 High quality sequence stop: 529.
 Location/Qualifiers

FEATURES

SOURCE

1. 529
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_plate="642 COL=9 ROW=J"
 /clone_lib="RPCI-11 Human Male BAC Library"
 /sex="male"
 /note="Vector: pBAC3.6; Site: 1: EcoRI; Site: 2: EcoRI;
 Male blood DNA was isolated from one randomly chosen donor
 and partially digested with a combination of EcoRI and
 EcoRI Methylase. Size selected DNA was cloned into the
 pBAC3.6 vector at EcoRI sites"
 145 a 99 c 109 g 165 t 11 others

BASE COUNT

ORIGIN

Query Match 79.1%; Score 17.4; DB 13: Length 529;
 Best Local Similarity 94.7%; Pred. No. 4.8e+02;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 ttggtcacagaatttgat 21
 |||||

Db 403 TTGGTCACAAGATTGGAT 421

RESULT 13

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AM695022 646 bp mRNA EST 21-DEC-2000
 N082E04ST1P1034 Developing stem Medicago truncatula cDNA clone
 N082E04ST 5', mRNA sequence.

ACCESSION AM695022
 VERSION AM695022.2 GI:11957200
 KEYWORDS EST.

SOURCE

ORGANISM

Medicago truncatula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;

Medicago.
 1 (bases 1 to 646)
 He,X.-Z., Shadle,G., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell
 C.J., Flores,H.R., Inman,J.T., Weller,J.W., May,G.D. and Dixon

,R.A.
 Expressed Sequence Tags from the Samuel Roberts Noble Foundation
 Medicago truncatula stem library

Unpublished (2000)
 On Apr 14, 2000 this sequence version replaced gi:7569784.

TITLE

JOURNAL

COMMENT

Contact: Dixon RA
 Plant Biology Division
 The Samuel Roberts Noble Foundation
 2510 Sam Noble Parkway, Ardmore, OK 73402, USA
 Tel: 580 221 7302
 Fax: 580 221 7380
 Email: radixon@noble.org

Insert Length: 674 Std Error: 0.00
 Plate: 082 row: E column: 04
 Seq primer: TCACACAGGAACAGCTATGAC.

FEATURES

Source

1. 646
 /organism="Medicago truncatula"
 /db_xref="taxon:3880"
 /clone="NF082E04ST"
 /clone_lib="Developing stem"
 /tissue_type="stem"
 /dev_stage="Pooled developmental"
 /note="Vector: Lambda Zap; Contains a mixture of
 intermodal stem segments"

BASE COUNT

ORIGIN

180 a 142 c 129 g 195 t

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 05:42:56 : Search time 7962.77 Seconds
(Without alignments)
15.407 Million cell updates/sec

Title: US-09-380-826a-5
Perfect score: 7
Sequence: 1 ttgata 7

Scoring table: IDENTITY-NUC
Gapop 10.0 , Gapext 1.0

Searched: 17159718 seqs, 8763200856 residues

Total number of hits satisfying chosen parameters: 34319436

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_NA_Main:*

1:	/cgn2_6/prodata/2/pna/PCRUS_COMB.seq:*
2:	/cgn2_6/prodata/2/pna/US06_COMB.seq:*
3:	/cgn2_6/prodata/2/pna/US07_COMB.seq:*
4:	/cgn2_6/prodata/2/pna/US080_COMB.seq:*
5:	/cgn2_6/prodata/2/pna/US081_COMB.seq:*
6:	/cgn2_6/prodata/2/pna/US082_COMB.seq:*
7:	/cgn2_6/prodata/2/pna/US083_COMB.seq:*
8:	/cgn2_6/prodata/2/pna/US084_COMB.seq:*
9:	/cgn2_6/prodata/2/pna/US085_COMB.seq:*
10:	/cgn2_6/prodata/2/pna/US086_COMB.seq:*
11:	/cgn2_6/prodata/2/pna/US087_COMB.seq:*
12:	/cgn2_6/prodata/2/pna/US088_COMB.seq:*
13:	/cgn2_6/prodata/2/pna/US089_COMB.seq:*
14:	/cgn2_6/prodata/2/pna/US090_COMB.seq:*
15:	/cgn2_6/prodata/2/pna/US091_COMB.seq:*
16:	/cgn2_6/prodata/2/pna/US092_COMB.seq:*
17:	/cgn2_6/prodata/2/pna/US093_COMB.seq:*
18:	/cgn2_6/prodata/2/pna/US094_COMB.seq:*
19:	/cgn2_6/prodata/2/pna/US095A_COMB.seq:*
20:	/cgn2_6/prodata/2/pna/US095B_COMB.seq:*
21:	/cgn2_6/prodata/2/pna/US095C_COMB.seq:*
22:	/cgn2_6/prodata/2/pna/US096A_COMB.seq:*
23:	/cgn2_6/prodata/2/pna/US096B_COMB.seq:*
24:	/cgn2_6/prodata/2/pna/US096C_COMB.seq:*
25:	/cgn2_6/prodata/2/pna/US096D_COMB.seq:*
26:	/cgn2_6/prodata/2/pna/US096E_COMB.seq:*
27:	/cgn2_6/prodata/2/pna/US096F_COMB.seq:*
28:	/cgn2_6/prodata/2/pna/US097A_COMB.seq:*
29:	/cgn2_6/prodata/2/pna/US097B_COMB.seq:*
30:	/cgn2_6/prodata/2/pna/US097C_COMB.seq:*
31:	/cgn2_6/prodata/2/pna/US098_COMB.seq:*
32:	/cgn2_6/prodata/2/pna/US099_COMB.seq:*
33:	/cgn2_6/prodata/2/pna/US6000_COMB.seq:*
34:	/cgn2_6/prodata/2/pna/US6001_COMB.seq:*
35:	/cgn2_6/prodata/2/pna/US6002_COMB.seq:*
36:	/cgn2_6/prodata/2/pna/US6003_COMB.seq:*
37:	/cgn2_6/prodata/2/pna/US6004_COMB.seq:*
38:	/cgn2_6/prodata/2/pna/US6005_COMB.seq:*
39:	/cgn2_6/prodata/2/pna/US6006_COMB.seq:*
40:	/cgn2_6/prodata/2/pna/US6007_COMB.seq:*
41:	/cgn2_6/prodata/2/pna/US6008_COMB.seq:*
42:	/cgn2_6/prodata/2/pna/US6009_COMB.seq:*
43:	/cgn2_6/prodata/2/pna/US6010_COMB.seq:*

Result				Query				SUMMARIES			
No.	Score	Match	Length	ID	Description	Pred. No.	Score	Match	Length	ID	Description
1	7	100.0	7	17	US-09-380-826a-5	44:	7	100.0	7	17	/cgn2_6/prodata/2/pna/US6011_COMB.seq:*
2	7	100.0	10	17	PCT-US99-13800-1005	45:	7	100.0	10	17	/cgn2_6/prodata/2/pna/US6012_COMB.seq:*
3	7	100.0	10	17	US-09-335-032-1287	46:	7	100.0	10	17	/cgn2_6/prodata/2/pna/US6013_COMB.seq:*
4	7	100.0	10	17	US-09-335-032-1793	47:	7	100.0	10	17	/cgn2_6/prodata/2/pna/US6014_COMB.seq:*
5	7	100.0	10	17	US-09-335-032-4831	48:	7	100.0	10	17	/cgn2_6/prodata/2/pna/US6015_COMB.seq:*
6	7	100.0	10	17	US-09-335-032-6568	49:	7	100.0	10	17	/cgn2_6/prodata/2/pna/US6016_COMB.seq:*
7	7	100.0	10	17	US-09-335-032-7108	50:	7	100.0	10	17	/cgn2_6/prodata/2/pna/US6017_COMB.seq:*
8	7	100.0	10	17	US-09-335-032-7153	51:	7	100.0	10	17	/cgn2_6/prodata/2/pna/US6018_COMB.seq:*
9	7	100.0	10	17	US-09-335-032-7153	52:	7	100.0	10	17	/cgn2_6/prodata/2/pna/US6019_COMB.seq:*
10	7	100.0	10	17	US-09-335-032-7517	53:	7	100.0	10	17	/cgn2_6/prodata/2/pna/US6020_COMB.seq:*
11	7	100.0	10	17	US-09-335-032-7816	54:	7	100.0	10	17	/cgn2_6/prodata/2/pna/US6021_COMB.seq:*
12	7	100.0	10	17	US-09-335-032-8062	55:	7	100.0	10	17	/cgn2_6/prodata/2/pna/US6022_COMB.seq:*
13	7	100.0	10	17	US-09-335-032-11663	56:	7	100.0	10	17	/cgn2_6/prodata/2/pna/US6023_COMB.seq:*
14	7	100.0	10	17	US-09-335-032-11665	57:	7	100.0	10	17	/cgn2_6/prodata/2/pna/US6024_COMB.seq:*
15	7	100.0	10	17	US-09-335-032-11665	58:	7	100.0	10	17	/cgn2_6/prodata/2/pna/US6025_COMB.seq:*
16	7	100.0	10	17	US-09-335-032-11665	59:	7	100.0	10	17	/cgn2_6/prodata/2/pna/US6026_COMB.seq:*
17	7	100.0	10	17	US-09-335-032-12098	60:	7	100.0	10	17	/cgn2_6/prodata/2/pna/US6027_COMB.seq:*
18	7	100.0	10	17	US-09-335-032-12170	61:	7	100.0	10	17	/cgn2_6/prodata/2/pna/US6028_COMB.seq:*
19	7	100.0	10	17	US-09-335-032-12170	62:	7	100.0	10	17	/cgn2_6/prodata/2/pna/US6029_COMB.seq:*
20	7	100.0	10	17	US-09-335-032-12170	63:	7	100.0	10	17	/cgn2_6/prodata/2/pna/US6030_COMB.seq:*
21	7	100.0	10	17	US-09-335-032-12170	64:	7	100.0	10	17	/cgn2_6/prodata/2/pna/US6031_COMB.seq:*
22	7	100.0	10	17	US-09-335-032-12170	65:	7	100.0	10	17	/cgn2_6/prodata/2/pna/US6032_COMB.seq:*
23	7	100.0	10	17	US-09-335-032-12170						
24	7	100.0	10	17	US-09-335-032-12170						
25	7	100.0	10	17	US-09-335-032-12170						
26	7	100.0	10	17	US-09-335-032-12170						
27	7	100.0	10	17	US-09-335-032-12170						
28	7	100.0	10	17	US-09-335-032-12170						
29	7	100.0	10	17	US-09-335-032-12170						
30	7	100.0	10	17	US-09-335-032-12170						
31	7	100.0	10	17	US-09-335-032-12170						
32	7	100.0	10	17	US-09-335-032-12170						
33	7	100.0	10	17	US-09-335-032-12170						
34	7	100.0	10	17	US-09-335-032-12170						
35	7	100.0	10	17	US-09-335-032-12170						
36	7	100.0	10	17	US-09-335-032-12170						
37	7	100.0	10	17	US-09-335-032-12170						
38	7	100.0	10	17	US-09-335-032-12170						
39	7	100.0	10	17	US-09-335-032-12170						
40	7	100.0	10	17	US-09-335-032-12170						
41	7	100.0	10	17	US-09-335-032-12170						
42	7	100.0	10	17	US-09-335-032-12170						
43	7	100.0	10	17	US-09-335-032-12170						

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

c 42 7 100.0 15 9 US-08-584-040-8444
c 43 7 100.0 15 9 US-08-585-684-113
c 44 7 100.0 15 9 US-08-585-684-114
c 45 7 100.0 15 28 US-09-708-690-4100

Sequence 8444, Ap
Sequence 113, App
Sequence 114, App
Sequence 4100, Ap

ALIGNMENTS

RESULT 1

US-09-380-826a-5
Sequence 5, Application US/09380826A
GENERAL INFORMATION:
APPLICANT: Chappel, Rod
TITLE OF INVENTION: LEPTOSPIRA PATHOGENS
FILE REFERENCE: DAVIE79.001APC
CURRENT APPLICATION NUMBER: US/09/380,826A
CURRENT FILING DATE: 1999-11-22
PRIOR APPLICATION NUMBER: PCT/AU98/00145
PRIOR FILING DATE: 1998-03-06
PRIOR APPLICATION NUMBER: AU P05494/97
PRIOR FILING DATE: 1997-05-07
NUMBER OF SEQ ID NOS: 26
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5
LENGTH: 7
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Primer for L. fainei.
US-09-380-826a-5

Query Match 100.0%; Score 7; DB 17; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.5e+09;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ttgata 7
Db 1 ttgata 7

RESULT 2

PCT-US99-13800-1005/c
Sequence 1005, Application PC/RTUS9913800
GENERAL INFORMATION:
APPLICANT: Genzyme Corporation
APPLICANT: Roberts, Bruce L.
APPLICANT: Shankara, Srihivas
TITLE OF INVENTION: PREPARATION AND USE OF SUPERIOR VACCINES
FILE REFERENCE: 68126881206940
CURRENT APPLICATION NUMBER: PCT/US99/13800
CURRENT FILING DATE: 1999-06-18
EARLIER APPLICATION NUMBER: 60/090,039
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,040
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,041
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/089,853
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/089,997
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,079
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,035
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/089,993
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/089,992
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,072
EARLIER FILING DATE: 1998-06-19

EARLIER APPLICATION NUMBER: 60/089,878
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/089,991
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,000
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,048
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/089,999
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,043
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,042
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,036
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,044
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/089,844
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,080
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/089,833
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/089,994
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,077
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,078
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,047
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,076
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,045
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/111,715
EARLIER FILING DATE: 1998-12-08
NUMBER OF SEQ ID NOS: 2138
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1005
LENGTH: 10
TYPE: DNA
ORGANISM: Homo sapiens
PCT-US99-13800-1005

Query Match 100.0%; Score 7; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ttgata 7
Db 9 TTGCATA 3

RESULT 3
US-09-335-032-1287/c
Sequence 1287, Application US/09335032
GENERAL INFORMATION:
APPLICANT: Velculescu, Victor
APPLICANT: Vogelstein, Bert
APPLICANT: Kinzler, Kenneth
TITLE OF INVENTION: Characterization of the Yeast
FILE REFERENCE: 01107.78572
CURRENT APPLICATION NUMBER: US/09/335,032
CURRENT FILING DATE: 1999-06-16
PRIOR APPLICATION NUMBER: US 60/035,917
PRIOR FILING DATE: 1997-01-23
PRIOR APPLICATION NUMBER: US 09/012,031
PRIOR FILING DATE: 1998-01-22
NUMBER OF SEQ ID NOS: 12219

```
SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1287
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-335-032-1287
```

```
Query Match          100.0%; Score 7; DB 17; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 ttgata 7
        |||||
Db      8 TTTGATA 2
```

```
RESULT 4
US-09-335-032-1793/c
; Sequence 1793, Application US/09335032
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the Yeast
; FILE REFERENCE: 01107.78572
; CURRENT APPLICATION NUMBER: US/09/335.032
; CURRENT FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1793
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-335-032-1793
```

```
Query Match          100.0%; Score 7; DB 17; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 ttgata 7
        |||||
Db      8 TTTGATA 2
```

```
RESULT 5
US-09-335-032-4831/c
; Sequence 4831, Application US/09335032
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the Yeast
; FILE REFERENCE: 01107.78572
; CURRENT APPLICATION NUMBER: US/09/335.032
; CURRENT FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4831
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
```

```
US-09-335-032-4831
```

```
Query Match          100.0%; Score 7; DB 17; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 ttgata 7
        |||||
Db      7 TTTGATA 1
```

```
RESULT 6
US-09-335-032-6568/c
; Sequence 6568, Application US/09335032
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the Yeast
; FILE REFERENCE: 01107.78572
; CURRENT APPLICATION NUMBER: US/09/335.032
; CURRENT FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6568
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-335-032-6568
```

```
Query Match          100.0%; Score 7; DB 17; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 ttgata 7
        |||||
Db      10 TTTGATA 4
```

```
RESULT 7
US-09-335-032-7108/c
; Sequence 7108, Application US/09335032
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the Yeast
; FILE REFERENCE: 01107.78572
; CURRENT APPLICATION NUMBER: US/09/335.032
; CURRENT FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7108
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-335-032-7108
```

```
Query Match          100.0%; Score 7; DB 17; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.3e+05;
```

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttgata 7
|||||
Db 10 TTTGATA 4

RESULT 8
US-09-335-032-7128
; Sequence 7128, Application US/09335032
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the Yeast
; FILE REFERENCE: 01107.78572
; CURRENT APPLICATION NUMBER: US/09/335,032
; CURRENT FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7128
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-335-032-7128

Query Match 100.0%; Score 7; DB 17; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttgata 7
|||||
Db 2 ttgata 8

RESULT 9
US-09-335-032-7153
; Sequence 7153, Application US/09335032
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the Yeast
; FILE REFERENCE: 01107.78572
; CURRENT APPLICATION NUMBER: US/09/335,032
; CURRENT FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7153
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-335-032-7153

Query Match 100.0%; Score 7; DB 17; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttgata 7
|||||
Db 3 ttgata 9

RESULT 10
US-09-335-032-7517
; Sequence 7517, Application US/09335032
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the Yeast
; FILE REFERENCE: 01107.78572
; CURRENT APPLICATION NUMBER: US/09/335,032
; CURRENT FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7517
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-335-032-7517

Query Match 100.0%; Score 7; DB 17; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttgata 7
|||||
Db 3 ttgata 9

RESULT 11
US-09-335-032-7816
; Sequence 7816, Application US/09335032
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the Yeast
; FILE REFERENCE: 01107.78572
; CURRENT APPLICATION NUMBER: US/09/335,032
; CURRENT FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7816
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-335-032-7816

Query Match 100.0%; Score 7; DB 17; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttgata 7
|||||
Db 3 ttgata 9

RESULT 12
US-09-335-032-8062/c
; Sequence 8062, Application US/09335032

```
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the Yeast
; FILE REFERENCE: 01107.78572
; CURRENT APPLICATION NUMBER: US/09/335,032
; CURRENT FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8062
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-335-032-8062
```

```
Query Match          100.0%; Score 7; DB 17; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 ttgata 7
    |||||
DB 8 TTTGATA 2
```

```
RESULT 13
US-09-335-032-10702/c
; Sequence 10702, Application US/09335032
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the Yeast
; FILE REFERENCE: 01107.78572
; CURRENT APPLICATION NUMBER: US/09/335,032
; CURRENT FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10702
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-335-032-10702
```

```
Query Match          100.0%; Score 7; DB 17; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 ttgata 7
    |||||
DB 7 TTTGATA 1
```

```
RESULT 14
US-09-335-032-11663/c
; Sequence 11663, Application US/09335032
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the Yeast
```

```
; TITLE OF INVENTION: Transcriptome
; FILE REFERENCE: 01107.78572
; CURRENT APPLICATION NUMBER: US/09/335,032
; CURRENT FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11663
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-335-032-11663
```

```
Query Match          100.0%; Score 7; DB 17; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 ttgata 7
    |||||
DB 8 TTTGATA 2
```

```
RESULT 15
US-09-335-032-11664/c
; Sequence 11664, Application US/09335032
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the Yeast
; FILE REFERENCE: 01107.78572
; CURRENT APPLICATION NUMBER: US/09/335,032
; CURRENT FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11664
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-335-032-11664
```

```
Query Match          100.0%; Score 7; DB 17; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 ttgata 7
    |||||
DB 8 TTTGATA 2
```

```
Search completed: December 15, 2001, 05:42:57
Job time: 15822 sec
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 05:58:20 : Search time 474.83 Seconds
(without alignments)
21.200 Million cell updates/sec

Title: US-09-380-826a-5
Perfect score: 7
Sequence: 1 ttgata 7

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 973175 seqs, 719035916 residues

Total number of hits satisfying chosen parameters: 1946350

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_New:*

1: /cgn2_6/ptodata/1/pna/PCN_NEW_COMB.seq:*
2: /cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq:*
3: /cgn2_6/ptodata/1/pna/US07_NEW_COMB.seq:*
4: /cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq:*
5: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:*
6: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq1:*
7: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:*
8: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	7	100.0	15	6	US-09-685-664B-4100 Sequence 4100, Ap
C 2	7	100.0	15	6	US-09-870-161-4100 Sequence 4100, Ap
C 3	7	100.0	16	6	US-09-670-607-1266 Sequence 1266, Ap
C 4	7	100.0	16	6	US-09-870-161-5694 Sequence 5694, Ap
C 5	7	100.0	16	6	US-09-870-161-5978 Sequence 5978, Ap
C 6	7	100.0	17	6	US-09-685-664B-116 Sequence 116, App
C 7	7	100.0	17	6	US-09-685-664B-117 Sequence 117, App
C 8	7	100.0	17	6	US-09-685-664B-118 Sequence 118, App
C 9	7	100.0	17	6	US-09-685-664B-882 Sequence 882, App
C 10	7	100.0	17	6	US-09-685-664B-883 Sequence 883, App
C 11	7	100.0	17	6	US-09-685-664B-2417 Sequence 2417, Ap
C 12	7	100.0	17	6	US-09-685-664B-2418 Sequence 2418, Ap
C 13	7	100.0	17	6	US-09-685-664B-3108 Sequence 3108, Ap
C 14	7	100.0	17	6	US-09-685-664B-3109 Sequence 3109, Ap
C 15	7	100.0	17	6	US-09-685-664B-3110 Sequence 3110, Ap
C 16	7	100.0	17	6	US-09-685-664B-3196 Sequence 3196, Ap
C 17	7	100.0	17	6	US-09-685-664B-3197 Sequence 3197, Ap
C 18	7	100.0	17	6	US-09-685-664B-3198 Sequence 3198, Ap
C 19	7	100.0	17	6	US-09-670-607-34 Sequence 34, App1
C 20	7	100.0	17	6	US-09-670-607-35 Sequence 35, App1
C 21	7	100.0	17	6	US-09-670-607-36 Sequence 36, App1
C 22	7	100.0	17	6	US-09-670-607-37 Sequence 37, App1
C 23	7	100.0	17	6	US-09-670-607-273 Sequence 273, App
C 24	7	100.0	17	6	US-09-670-607-274 Sequence 274, App
C 25	7	100.0	17	6	US-09-670-607-275 Sequence 275, App

C 26	7	100.0	17	6	US-09-670-607-276 Sequence 276, App
C 27	7	100.0	17	6	US-09-670-607-971 Sequence 971, App
C 28	7	100.0	17	6	US-09-670-607-972 Sequence 972, App
C 29	7	100.0	17	6	US-09-670-607-1467 Sequence 1467, App
C 30	7	100.0	17	6	US-09-670-607-1468 Sequence 1468, App
C 31	7	100.0	17	6	US-09-870-161-116 Sequence 116, App
C 32	7	100.0	17	6	US-09-870-161-117 Sequence 117, App
C 33	7	100.0	17	6	US-09-870-161-118 Sequence 118, App
C 34	7	100.0	17	6	US-09-870-161-882 Sequence 882, App
C 35	7	100.0	17	6	US-09-870-161-883 Sequence 883, App
C 36	7	100.0	17	6	US-09-870-161-2417 Sequence 2417, App
C 37	7	100.0	17	6	US-09-870-161-2418 Sequence 2418, App
C 38	7	100.0	17	6	US-09-870-161-3108 Sequence 3108, App
C 39	7	100.0	17	6	US-09-870-161-3109 Sequence 3109, App
C 40	7	100.0	17	6	US-09-870-161-3110 Sequence 3110, App
C 41	7	100.0	17	6	US-09-870-161-3196 Sequence 3196, App
C 42	7	100.0	17	6	US-09-870-161-3197 Sequence 3197, App
C 43	7	100.0	17	6	US-09-870-161-3198 Sequence 3198, App
C 44	7	100.0	17	6	US-09-870-161-4320 Sequence 4320, App
C 45	7	100.0	17	6	US-09-870-161-4321 Sequence 4321, App

ALIGNMENTS

```
RESULT 1
US-09-685-664B-4100/c
: Sequence 4100, Application US/09685664B
: GENERAL INFORMATION:
: APPLICANT: Ribozyme Pharmaceuticals, Inc.
: APPLICANT: Pavco, Pam
: APPLICANT: McSwigen, Jim
: APPLICANT: Stinchcomb, Dan
: APPLICANT: Escobedo, Jaime
: TITLE OF INVENTION: Method and Reagent for Treatment of Diseases or Conditions Rel
: FILE REFERENCE: MBH00-876-K (400/021)
: CURRENT APPLICATION NUMBER: US/09/685,664B
: PRIOR FILING DATE: 2000-10-10
: PRIOR APPLICATION NUMBER: US 60/005,974
: PRIOR FILING DATE: 1995-10-26
: PRIOR APPLICATION NUMBER: US 08/584,040
: PRIOR FILING DATE: 1996-01-08
: PRIOR APPLICATION NUMBER: US 09/371,772
: PRIOR FILING DATE: 1999-08-10
: NUMBER OF SEQ ID NOS: 8231
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 4100
: LENGTH: 15
: TYPE: RNA
: ORGANISM: Homo sapiens
US-09-685-664B-4100

Query Match 100.0%; Score 7; DB 6; Length 15;
Best Local Similarity 100.0%; Pred. No. 7.1e+04;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ttgata 7
Db 13 TTTGATA 7

RESULT 2
US-09-870-161-4100/c
: Sequence 4100, Application US/09870161
: GENERAL INFORMATION:
: APPLICANT: Ribozyme Pharmaceuticals, Inc.
: APPLICANT: Pavco, Pam
: APPLICANT: McSwigen, Jim
: APPLICANT: Stinchcomb, Dan
: APPLICANT: Escobedo, Jaime
: TITLE OF INVENTION: Method and Reagent for the Treatment of Diseases or Conditions
```

```
;; TITLE OF INVENTION: Levels of Vascular Endothelial Growth Factor Receptor
;; FILE REFERENCE: MBHB00-876-M (400/026)
;; CURRENT APPLICATION NUMBER: US/09/870,161
;; CURRENT FILING DATE: 2001-08-27
;; NUMBER OF SEQ ID NOS: 20821
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 4100
;; LENGTH: 15
;; TYPE: RNA
;; ORGANISM: Homo sapiens
US-09-870-161-4100
```

```
Query Match          100.0%; Score 7; DB 6; Length 15;
Best Local Similarity 100.0%; Pred. No. 7.1e+04;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 ttgata 7
        |||||
Db      13 TTGATA 7
```

```
RESULT 3
US-09-670-607-1266
; Sequence 1266, Application US/09670607
; GENERAL INFORMATION:
; APPLICANT: Ribozyne Pharmaceuticals, Inc.
; APPLICANT: Usman, Nassim
; APPLICANT: MCSwigen, Jim
; TITLE OF INVENTION: Method and Reagent for the Inhibition of Protein-Tyrosine Phosphatase
; FILE REFERENCE: MBHB00-836-A (237/194)
; CURRENT APPLICATION NUMBER: US/09/670,607
; CURRENT FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 4262
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1266
; LENGTH: 16
; TYPE: RNA
; ORGANISM: Homo sapiens
US-09-670-607-1266
```

```
Query Match          100.0%; Score 7; DB 6; Length 16;
Best Local Similarity 42.9%; Pred. No. 7.1e+04;
Matches 3; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 ttgata 7
        ::|||
Db      8 uuugaaa 14
```

```
RESULT 4
US-09-870-161-5694/C
; Sequence 5694, Application US/09870161
; GENERAL INFORMATION:
; APPLICANT: Ribozyne Pharmaceuticals, Inc.
; APPLICANT: Pavco, Pam
; APPLICANT: MCSwigen, Jim
; APPLICANT: Stinchcomb, Dan
; APPLICANT: Escobedo, Jaime
; TITLE OF INVENTION: Method and Reagent for the Treatment of Diseases or Conditions Related to Vascular Endothelial Growth Factor Receptor
; FILE REFERENCE: MBHB00-876-M (400/026)
; CURRENT APPLICATION NUMBER: US/09/870,161
; CURRENT FILING DATE: 2001-08-27
; NUMBER OF SEQ ID NOS: 20821
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5694
; LENGTH: 16
; TYPE: RNA
; ORGANISM: Homo sapiens
US-09-870-161-5694
```

```
Query Match          100.0%; Score 7; DB 6; Length 16;
Best Local Similarity 100.0%; Pred. No. 7.1e+04;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 ttgata 7
        |||||
Db      9 TTGATA 3
```

```
RESULT 5
US-09-870-161-5978/C
; Sequence 5978, Application US/09870161
; GENERAL INFORMATION:
; APPLICANT: Ribozyne Pharmaceuticals, Inc.
; APPLICANT: Pavco, Pam
; APPLICANT: MCSwigen, Jim
; APPLICANT: Stinchcomb, Dan
; APPLICANT: Escobedo, Jaime
; TITLE OF INVENTION: Method and Reagent for the Treatment of Diseases or Conditions Related to Vascular Endothelial Growth Factor Receptor
; FILE REFERENCE: MBHB00-876-M (400/026)
; CURRENT APPLICATION NUMBER: US/09/870,161
; CURRENT FILING DATE: 2001-08-27
; NUMBER OF SEQ ID NOS: 20821
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5978
; LENGTH: 16
; TYPE: RNA
; ORGANISM: Homo sapiens
US-09-870-161-5978
```

```
Query Match          100.0%; Score 7; DB 6; Length 16;
Best Local Similarity 100.0%; Pred. No. 7.1e+04;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 ttgata 7
        |||||
Db      8 TTGATA 2
```

```
RESULT 6
US-09-685-664B-116/C
; Sequence 116, Application US/09685664B
; GENERAL INFORMATION:
; APPLICANT: Ribozyne Pharmaceuticals, Inc.
; APPLICANT: Pavco, Pam
; APPLICANT: MCSwigen, Jim
; APPLICANT: Stinchcomb, Dan
; APPLICANT: Escobedo, Jaime
; TITLE OF INVENTION: Method and Reagent for Treatment of Diseases or Conditions Related to Vascular Endothelial Growth Factor Receptor
; FILE REFERENCE: MBHB00-876-K (400/021)
; CURRENT APPLICATION NUMBER: US/09/685,664B
; CURRENT FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: US 60/005,974
; PRIOR FILING DATE: 1995-10-26
; PRIOR APPLICATION NUMBER: US 08/584,040
; PRIOR FILING DATE: 1996-01-08
; PRIOR APPLICATION NUMBER: US 09/371,772
; PRIOR FILING DATE: 1999-08-10
; NUMBER OF SEQ ID NOS: 8231
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 116
; LENGTH: 17
; TYPE: RNA
; ORGANISM: Homo sapiens
US-09-685-664B-116
```

```
Query Match          100.0%; Score 7; DB 6; Length 17;
```

Best Local Similarity 100.0%; Pred. No. 7.1e+04;
Matches 7; Conservative 0; Mismatches 0; Indels 0;

QY 1 ttgata 7
|||||||
DB 17 TTTGATA 11

RESULT 7
US-09-685-664B-117/C

; Sequence 117, Application US/09685664B
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals, Inc.
; APPLICANT: Pavco, Pam
; APPLICANT: McSwiggen, Jim
; APPLICANT: Stinchcomb, Dan
; APPLICANT: Escobedo, Jaime
; TITLE OF INVENTION: Method and Reagent for Treatment of Diseases or Conditions Related to
; TITLE OF INVENTION: Levels of Vascular Endothelial Growth Factor Receptor
; FILE REFERENCE: MBH00-876-K (400/021)
; CURRENT APPLICATION NUMBER: US/09/685,664B
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: US 60/005,974
; PRIOR FILING DATE: 1995-10-26
; PRIOR APPLICATION NUMBER: US 08/584,040
; PRIOR FILING DATE: 1996-01-08
; PRIOR APPLICATION NUMBER: US 09/371,772
; PRIOR FILING DATE: 1999-08-10
; NUMBER OF SEQ ID NOS: 8231
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 117
; LENGTH: 17
; TYPE: RNA
; ORGANISM: Homo sapiens
US-09-685-664B-117

Query Match 100.0%; Score 7; DB 6; Length 17;
Best Local Similarity 100.0%; Pred. No. 7.1e+04;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ttgata 7
|||||||
DB 14 TTTGATA 8

RESULT 8

US-09-685-664B-118/C
; Sequence 118, Application US/09685664B
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals, Inc.
; APPLICANT: Pavco, Pam
; APPLICANT: McSwiggen, Jim
; APPLICANT: Stinchcomb, Dan
; APPLICANT: Escobedo, Jaime
; TITLE OF INVENTION: Method and Reagent for Treatment of Diseases or Conditions Related to
; TITLE OF INVENTION: Levels of Vascular Endothelial Growth Factor Receptor
; FILE REFERENCE: MBH00-876-K (400/021)
; CURRENT APPLICATION NUMBER: US/09/685,664B
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: US 60/005,974
; PRIOR FILING DATE: 1995-10-26
; PRIOR APPLICATION NUMBER: US 08/584,040
; PRIOR FILING DATE: 1996-01-08
; PRIOR APPLICATION NUMBER: US 09/371,772
; PRIOR FILING DATE: 1999-08-10
; NUMBER OF SEQ ID NOS: 8231
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 118
; LENGTH: 17
; TYPE: RNA
; ORGANISM: Homo sapiens
US-09-685-664B-118

Query Match 100.0%; Score 7; DB 6; Length 17;
Best Local Similarity 100.0%; Pred. No. 7.1e+04;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ttgata 7
|||||||
DB 12 TTTGATA 6

RESULT 9

US-09-685-664B-882/C
; Sequence 882, Application US/09685664B
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals, Inc.
; APPLICANT: Pavco, Pam
; APPLICANT: McSwiggen, Jim
; APPLICANT: Stinchcomb, Dan
; APPLICANT: Escobedo, Jaime
; TITLE OF INVENTION: Method and Reagent for Treatment of Diseases or Conditions Related to
; TITLE OF INVENTION: Levels of Vascular Endothelial Growth Factor Receptor
; FILE REFERENCE: MBH00-876-K (400/021)
; CURRENT APPLICATION NUMBER: US/09/685,664B
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: US 60/005,974
; PRIOR FILING DATE: 1995-10-26
; PRIOR APPLICATION NUMBER: US 08/584,040
; PRIOR FILING DATE: 1996-01-08
; PRIOR APPLICATION NUMBER: US 09/371,772
; PRIOR FILING DATE: 1999-08-10
; NUMBER OF SEQ ID NOS: 8231
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 882
; LENGTH: 17
; TYPE: RNA
; ORGANISM: Homo sapiens
US-09-685-664B-882

Query Match 100.0%; Score 7; DB 6; Length 17;
Best Local Similarity 100.0%; Pred. No. 7.1e+04;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ttgata 7
|||||||
DB 14 TTTGATA 8

RESULT 10

US-09-685-664B-883/C
; Sequence 883, Application US/09685664B
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals, Inc.
; APPLICANT: Pavco, Pam
; APPLICANT: McSwiggen, Jim
; APPLICANT: Stinchcomb, Dan
; APPLICANT: Escobedo, Jaime
; TITLE OF INVENTION: Method and Reagent for Treatment of Diseases or Conditions Related to
; TITLE OF INVENTION: Levels of Vascular Endothelial Growth Factor Receptor
; FILE REFERENCE: MBH00-876-K (400/021)
; CURRENT APPLICATION NUMBER: US/09/685,664B
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: US 60/005,974
; PRIOR FILING DATE: 1995-10-26
; PRIOR APPLICATION NUMBER: US 08/584,040
; PRIOR FILING DATE: 1996-01-08
; PRIOR APPLICATION NUMBER: US 09/371,772
; PRIOR FILING DATE: 1999-08-10
; NUMBER OF SEQ ID NOS: 8231
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 883
; LENGTH: 17

```

; TYPE: RNA
; ORGANISM: Homo sapiens
US-09-685-664B-883

```

```

Query Match          100.0%; Score 7; DB 6; Length 17;
Best Local Similarity 100.0%; Pred. No. 7.1e+04;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 ttgata 7
    |||||
Db 12 TTTGATA 6

```

```

RESULT 11
US-09-685-664B-2417/c
; Sequence 2417, Application US/09685664B
; GENERAL INFORMATION:
; APPLICANT: Ribozyne Pharmaceuticals, Inc.
; APPLICANT: Pavco, Pam
; APPLICANT: MCSwigen, Jim
; APPLICANT: Stinchcomb, Dan
; APPLICANT: Escobedo, Jaime
; TITLE OF INVENTION: Method and Reagent for Treatment of Diseases or Conditions Relate
; TITLE OF INVENTION: Levels of Vascular Endothelial Growth Factor Receptor
; FILE REFERENCE: MBH00-876-K (400/021)
; CURRENT APPLICATION NUMBER: US/09/685,664B
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: US 60/005,974
; PRIOR FILING DATE: 1995-10-26
; PRIOR APPLICATION NUMBER: US 08/584,040
; PRIOR FILING DATE: 1996-01-08
; PRIOR APPLICATION NUMBER: US 09/371,772
; PRIOR FILING DATE: 1999-08-10
; NUMBER OF SEQ ID NOS: 8231
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2417
; LENGTH: 17
; TYPE: RNA
; ORGANISM: Mus musculus
US-09-685-664B-2417

```

```

Query Match          100.0%; Score 7; DB 6; Length 17;
Best Local Similarity 100.0%; Pred. No. 7.1e+04;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 ttgata 7
    |||||
Db 14 TTTGATA 8

```

```

RESULT 12
US-09-685-664B-2418/c
; Sequence 2418, Application US/09685664B
; GENERAL INFORMATION:
; APPLICANT: Ribozyne Pharmaceuticals, Inc.
; APPLICANT: Pavco, Pam
; APPLICANT: MCSwigen, Jim
; APPLICANT: Stinchcomb, Dan
; APPLICANT: Escobedo, Jaime
; TITLE OF INVENTION: Method and Reagent for Treatment of Diseases or Conditions Relate
; TITLE OF INVENTION: Levels of Vascular Endothelial Growth Factor Receptor
; FILE REFERENCE: MBH00-876-K (400/021)
; CURRENT APPLICATION NUMBER: US/09/685,664B
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: US 60/005,974
; PRIOR FILING DATE: 1995-10-26
; PRIOR APPLICATION NUMBER: US 08/584,040
; PRIOR FILING DATE: 1996-01-08
; PRIOR APPLICATION NUMBER: US 09/371,772
; PRIOR FILING DATE: 1999-08-10
; NUMBER OF SEQ ID NOS: 8231

```

```

; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2418
; LENGTH: 17
; TYPE: RNA
; ORGANISM: Mus musculus
US-09-685-664B-2418

```

```

Query Match          100.0%; Score 7; DB 6; Length 17;
Best Local Similarity 100.0%; Pred. No. 7.1e+04;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 ttgata 7
    |||||
Db 12 TTTGATA 6

```

```

RESULT 13
US-09-685-664B-3108
; Sequence 3108, Application US/09685664B
; GENERAL INFORMATION:
; APPLICANT: Ribozyne Pharmaceuticals, Inc.
; APPLICANT: Pavco, Pam
; APPLICANT: MCSwigen, Jim
; APPLICANT: Stinchcomb, Dan
; APPLICANT: Escobedo, Jaime
; TITLE OF INVENTION: Method and Reagent for Treatment of Diseases or Conditions Rel
; TITLE OF INVENTION: Levels of Vascular Endothelial Growth Factor Receptor
; FILE REFERENCE: MBH00-876-K (400/021)
; CURRENT APPLICATION NUMBER: US/09/685,664B
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: US 60/005,974
; PRIOR FILING DATE: 1995-10-26
; PRIOR APPLICATION NUMBER: US 08/584,040
; PRIOR FILING DATE: 1996-01-08
; PRIOR APPLICATION NUMBER: US 09/371,772
; PRIOR FILING DATE: 1999-08-10
; NUMBER OF SEQ ID NOS: 8231
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3108
; LENGTH: 17
; TYPE: RNA
; ORGANISM: Mus musculus
US-09-685-664B-3108

```

```

Query Match          100.0%; Score 7; DB 6; Length 17;
Best Local Similarity 42.9%; Pred. No. 7.1e+04;
Matches 3; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 ttgata 7
    ::|||:|
Db 8 uuugaaua 14

```

```

RESULT 14
US-09-685-664B-3109
; Sequence 3109, Application US/09685664B
; GENERAL INFORMATION:
; APPLICANT: Ribozyne Pharmaceuticals, Inc.
; APPLICANT: Pavco, Pam
; APPLICANT: MCSwigen, Jim
; APPLICANT: Stinchcomb, Dan
; APPLICANT: Escobedo, Jaime
; TITLE OF INVENTION: Method and Reagent for Treatment of Diseases or Conditions Rel
; TITLE OF INVENTION: Levels of Vascular Endothelial Growth Factor Receptor
; FILE REFERENCE: MBH00-876-K (400/021)
; CURRENT APPLICATION NUMBER: US/09/685,664B
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: US 60/005,974
; PRIOR FILING DATE: 1995-10-26
; PRIOR APPLICATION NUMBER: US 08/584,040
; PRIOR FILING DATE: 1996-01-08

```

; PRIOR APPLICATION NUMBER: US 09/371,772
 ; PRIOR FILING DATE: 1999-08-10
 ; NUMBER OF SEQ ID NOS: 8231
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 3109
 ; LENGTH: 17
 ; TYPE: RNA
 ; ORGANISM: Mus musculus
 US-09-685-664B-3109

Query Match 100.0%; Score 7; DB 6; Length 17;
 Best Local Similarity 42.9%; Pred. No. 7.1e+04;
 Matches 3; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 ttgata 7
 : : : | : | : |
 Db 7 uuugaua 13

RESULT 15
 US-09-685-664B-3110
 ; Sequence 3110, Application US/09685664B
 ; GENERAL INFORMATION:
 ; APPLICANT: Ribozyne Pharmaceuticals, Inc.
 ; APPLICANT: Pavco, Pam
 ; APPLICANT: McSwigen, Jim
 ; APPLICANT: Stinchcomb, Dan
 ; APPLICANT: Escobedo, Jaime
 ; TITLE OF INVENTION: Method and Reagent for Treatment of Diseases or Conditions Relate
 ; FILE REFERENCE: MBHB00-876-K (400/021)
 ; CURRENT APPLICATION NUMBER: US/09/685,664B
 ; CURRENT FILING DATE: 2000-10-10
 ; PRIOR APPLICATION NUMBER: US 60/005,974
 ; PRIOR FILING DATE: 1995-10-26
 ; PRIOR APPLICATION NUMBER: US 08/584,040
 ; PRIOR FILING DATE: 1996-01-08
 ; PRIOR APPLICATION NUMBER: US 09/371,772
 ; PRIOR FILING DATE: 1999-08-10
 ; NUMBER OF SEQ ID NOS: 8231
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 3110
 ; LENGTH: 17
 ; TYPE: RNA
 ; ORGANISM: Mus musculus
 US-09-685-664B-3110

Query Match 100.0%; Score 7; DB 6; Length 17;
 Best Local Similarity 42.9%; Pred. No. 7.1e+04;
 Matches 3; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 ttgata 7
 : : : | : | : |
 Db 3 uuugaua 9

Search completed: December 15, 2001, 05:58:20
 Job time: 12263 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 05:42:57 ; Search time 7962.77 Seconds
(without alignments) 48.423 Million cell updates/sec

Title: US-09-380-826a-6
Perfect score: 22
Sequence: 1 tfgtggannnnnnnttgcata-22

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 17159718 seqs, 8763200856 residues

Total number of hits satisfying chosen parameters: 34319436

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NM_Main:*

```
1: /cgn2_6/ptodata/2/pna/PCYUS_COMB.seq:*
2: /cgn2_6/ptodata/2/pna/US06_COMB.seq:*
3: /cgn2_6/ptodata/2/pna/US07_COMB.seq:*
4: /cgn2_6/ptodata/2/pna/US08_COMB.seq:*
5: /cgn2_6/ptodata/2/pna/US081_COMB.seq:*
6: /cgn2_6/ptodata/2/pna/US082_COMB.seq:*
7: /cgn2_6/ptodata/2/pna/US083_COMB.seq:*
8: /cgn2_6/ptodata/2/pna/US084_COMB.seq:*
9: /cgn2_6/ptodata/2/pna/US085_COMB.seq:*
10: /cgn2_6/ptodata/2/pna/US086_COMB.seq:*
11: /cgn2_6/ptodata/2/pna/US087_COMB.seq:*
12: /cgn2_6/ptodata/2/pna/US088_COMB.seq:*
13: /cgn2_6/ptodata/2/pna/US089_COMB.seq:*
14: /cgn2_6/ptodata/2/pna/US090_COMB.seq:*
15: /cgn2_6/ptodata/2/pna/US091_COMB.seq:*
16: /cgn2_6/ptodata/2/pna/US092_COMB.seq:*
17: /cgn2_6/ptodata/2/pna/US093_COMB.seq:*
18: /cgn2_6/ptodata/2/pna/US094_COMB.seq:*
19: /cgn2_6/ptodata/2/pna/US095_COMB.seq:*
20: /cgn2_6/ptodata/2/pna/US095A_COMB.seq:*
21: /cgn2_6/ptodata/2/pna/US095C_COMB.seq:*
22: /cgn2_6/ptodata/2/pna/US095D_COMB.seq:*
23: /cgn2_6/ptodata/2/pna/US096_COMB.seq:*
24: /cgn2_6/ptodata/2/pna/US096B_COMB.seq:*
25: /cgn2_6/ptodata/2/pna/US096C_COMB.seq:*
26: /cgn2_6/ptodata/2/pna/US096D_COMB.seq:*
27: /cgn2_6/ptodata/2/pna/US096E_COMB.seq:*
28: /cgn2_6/ptodata/2/pna/US097A_COMB.seq:*
29: /cgn2_6/ptodata/2/pna/US097B_COMB.seq:*
30: /cgn2_6/ptodata/2/pna/US097C_COMB.seq:*
31: /cgn2_6/ptodata/2/pna/US098_COMB.seq:*
32: /cgn2_6/ptodata/2/pna/US099_COMB.seq:*
33: /cgn2_6/ptodata/2/pna/US099A_COMB.seq:*
34: /cgn2_6/ptodata/2/pna/US099B_COMB.seq:*
35: /cgn2_6/ptodata/2/pna/US099C_COMB.seq:*
36: /cgn2_6/ptodata/2/pna/US099D_COMB.seq:*
37: /cgn2_6/ptodata/2/pna/US099E_COMB.seq:*
38: /cgn2_6/ptodata/2/pna/US099F_COMB.seq:*
39: /cgn2_6/ptodata/2/pna/US099G_COMB.seq:*
40: /cgn2_6/ptodata/2/pna/US099H_COMB.seq:*
41: /cgn2_6/ptodata/2/pna/US099I_COMB.seq:*
42: /cgn2_6/ptodata/2/pna/US099J_COMB.seq:*
43: /cgn2_6/ptodata/2/pna/US099K_COMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	14	63.6	22	17	US-09-380-826a-6
2	14	63.6	22	17	US-09-380-826a-6
3	14	63.6	22	17	US-09-380-826a-6
4	14	63.6	147	18	US-09-497-191-767
5	14	63.6	243	11	US-08-725-863-2005
6	14	63.6	243	51	US-60-184-813-200
7	14	63.6	256	45	US-60-127-461-5765
8	14	63.6	259	12	US-08-812-505-2536
9	14	63.6	259	13	US-08-978-620-2536
10	14	63.6	259	34	US-60-013-529-2536
11	14	63.6	261	18	US-09-411-077-732
12	14	63.6	261	21	US-09-540-213-52141
13	14	63.6	261	43	US-60-103-748-618
14	14	63.6	317	30	US-09-770-255-141
15	14	63.6	322	28	US-09-705-926-17256
16	14	63.6	326	25	US-09-654-617-181276
17	14	63.6	326	27	US-09-684-016-181276
18	14	63.6	333	24	US-09-637-086A-25736
19	14	63.6	333	25	US-09-654-617-216399
20	14	63.6	333	27	US-09-684-016-216399
21	14	63.6	353	19	US-09-521-640-131370
22	14	63.6	365	54	US-60-211-750-49225
23	14	63.6	385	17	US-09-371-805-2913
24	14	63.6	385	17	US-09-362-510-50353
25	14	63.6	393	17	US-09-362-510-50353
26	14	63.6	393	19	US-09-515-126-11383
27	14	63.6	393	22	US-09-577-409-11383
28	14	63.6	396	19	US-09-515-128-18844
29	14	63.6	398	17	US-09-362-510-50840
30	14	63.6	398	17	US-09-362-510A-50840
31	14	63.6	420	16	US-09-248-796-13484
32	14	63.6	420	42	US-60-096-409-13484
33	14	63.6	427	17	US-09-304-517A-90042
34	14	63.6	427	17	US-09-371-146A-90042
35	14	63.6	427	22	US-09-553-094-2550
36	14	63.6	427	25	US-09-654-617-292752
37	14	63.6	427	27	US-09-684-016-292752
38	14	63.6	427	46	US-60-130-180-1137
39	14	63.6	435	17	US-09-342-069-3899
40	14	63.6	435	17	US-09-342-069A-3899
41	14	63.6	435	17	US-09-342-069A-3899

```
c 42      14      63.6      456      53      US-60-207-458-38165      Sequence 38165, A
c 43      14      63.6      461      17      US-09-304-517A-85089      Sequence 85089, A
c 44      14      63.6      461      25      US-09-654-617-291806      Sequence 291806,
c 45      14      63.6      461      27      US-09-684-016-291806      Sequence 291806,
```

ALIGNMENTS

```
RESULT 1
US-09-380-826A-2
; Sequence 2, Application US/09380826A
; GENERAL INFORMATION:
; APPLICANT: Chappel, Rod
; TITLE OF INVENTION: LEPTOSPIRA PATHOGENS
; FILE REFERENCE: DAVIE79.001APC
; CURRENT APPLICATION NUMBER: US/09/380,826A
; PRIOR FILING DATE: 1999-11-22
; PRIOR APPLICATION NUMBER: PCT/AU98/00145
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: AU P05494/97
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Primer for L. fainei.
US-09-380-826A-2
```

```
Query Match      63.6%; Score 14; DB 17; Length 22;
Best Local Similarity 63.6%; Pred. No. 7.3e+02;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
```

```
QY 1 tgttgagannnnnnnttgata 22
|||||
Db 1 tgttgatcacagaattgata 22
```

```
RESULT 2
US-09-380-826A-6
; Sequence 6, Application US/09380826A
; GENERAL INFORMATION:
; APPLICANT: Chappel, Rod
; TITLE OF INVENTION: LEPTOSPIRA PATHOGENS
; FILE REFERENCE: DAVIE79.001APC
; CURRENT APPLICATION NUMBER: US/09/380,826A
; PRIOR FILING DATE: 1999-11-22
; PRIOR APPLICATION NUMBER: PCT/AU98/00145
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: AU P05494/97
; PRIOR FILING DATE: 1997-03-07
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(22)
; OTHER INFORMATION: n = A,T,C or G
; OTHER INFORMATION: Primer for L. fainei.
US-09-380-826A-6
```

```
Query Match      63.6%; Score 14; DB 17; Length 22;
Best Local Similarity 100.0%; Pred. No. 7.3e+02;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 tgttgagannnnnnnttgata 22
|||||
Db 1 tgttgatcacagaattgata 22
```

```
RESULT 3
US-09-380-826A-7
; Sequence 7, Application US/09380826A
; GENERAL INFORMATION:
; APPLICANT: Chappel, Rod
; TITLE OF INVENTION: LEPTOSPIRA PATHOGENS
; FILE REFERENCE: DAVIE79.001APC
; CURRENT APPLICATION NUMBER: US/09/380,826A
; PRIOR FILING DATE: 1999-11-22
; PRIOR APPLICATION NUMBER: PCT/AU98/00145
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: AU P05494/97
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Primer for L. fainei.
US-09-380-826A-7
```

```
Query Match      63.6%; Score 14; DB 17; Length 22;
Best Local Similarity 63.6%; Pred. No. 7.3e+02;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
```

```
QY 1 tgttgagannnnnnnttgata 22
|||||
Db 1 tgttgatcacagaattgata 22
```

```
RESULT 4
US-09-497-191-767/c
; Sequence 767, Application US/09497191
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai
; APPLICANT: CHEN, Xianfeng
; APPLICANT: SUBRAMANIAN, Gopalakrishnan
; APPLICANT: ZHENG, Liansheng
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS 84-89, 91-99
; FILE REFERENCE: 2750-694P
; CURRENT APPLICATION NUMBER: US/09/497,191
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 920
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 767
; LENGTH: 147
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: any n or xaa = unknown
US-09-497-191-767
```

```
Query Match      63.6%; Score 14; DB 18; Length 147;
Best Local Similarity 63.6%; Pred. No. 1.1e+03;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
```

```
QY 1 tgttgagannnnnnnttgata 22
|||||
Db 55 TGTGCACTTGACTTTTGATA 34
```

```
RESULT 5
US-08-725-863-2005
; Sequence 2005, Application US/08725863
```



```

: GENERAL INFORMATION:
: APPLICANT: Stuart, Susan G.
: APPLICANT: Gooding, Douglas H.
: APPLICANT: Lane, John C.
: APPLICANT: Deleage, Angelo M.
: APPLICANT: Shable, James I.
: TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
: TITLE OF INVENTION: HUMAN NEUTROPHILS
: NUMBER OF SEQUENCES: 3631
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
: STREET: 3174 PORTER DRIVE
: CITY: PALO ALTO
: STATE: CALIFORNIA
: COUNTRY: USA
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/725, 863
: FILING DATE: Herewith
: ATTORNEY/AGENT INFORMATION:
: NAME: CERRONE, MICHAEL C., Ph. D.
: REGISTRATION NUMBER: 39,132
: REFERENCE/DOCKET NUMBER: PD-0043 US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 855-0555
: TELEFAX: (415) 845-4166
: INFORMATION FOR SEQ ID NO: 2005:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 243 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: IMMEDIATE SOURCE:
: CLONE: 342263
: US-08-725-863-2005

Query Match          63.6%; Score 14; DB 11; Length 243;
Best Local Similarity 63.6%; Pred. No. 1.2e+03;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 tcttgagannnnnnnttgata 22
      |||||
Db 152 tcttgagatttcgtttgata 173

RESULT 6
US-60-184-813-200
: Sequence 200, Application US/60184813
: GENERAL INFORMATION:
: APPLICANT: Hodgson, David M.
: APPLICANT: Lincoln, Stephen E.
: APPLICANT: Russo, Frank D.
: APPLICANT: Spillo, Peter A.
: APPLICANT: Banville, Steve C.
: APPLICANT: Bratcher, Shawn R.
: APPLICANT: Dufour, Gerard E.
: APPLICANT: Cohen, Howard J.
: APPLICANT: Rosen, Bruce
: APPLICANT: Shah, Purvi
: APPLICANT: Chai, Michael S.
: APPLICANT: Hillman, Jennifer L.
: TITLE OF INVENTION: NUCLEIC ACID SYNTHESIS AND MODIFICATION ENZYMES
: FILE REFERENCE: PT-0121 P
: CURRENT APPLICATION NUMBER: US/60/184,813
: CURRENT FILING DATE: 2000-02-24
: NUMBER OF SEQ ID NOS: 905
: Jones, Anissa L.
: Yu, Jimmy Y.
: Greenwalt, Lila B.
: Panzer, Scott R.
: Roseberry, Ann M.
: Wright, Rachel J.
: Chen, Wensheng
: Liu, Tommy
: Yap, Pierre E.
: Stockdeher, Theresa K.
: Amshay, Stefan
: Fong, Willy Tuen
```

```

: SOFTWARE: PERL Program
: SEQ ID NO 200
: LENGTH: 243
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: OTHER INFORMATION: Incyte ID NO: 1050812.1
: FEATURE:
: NAME/KEY: unsure
: LOCATION: 130
: OTHER INFORMATION: a, t, c, g, or other
: US-60-184-813-200

Query Match          63.6%; Score 14; DB 51; Length 243;
Best Local Similarity 63.6%; Pred. No. 1.2e+03;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 tcttgagannnnnnnttgata 22
      |||||
Db 152 tcttgagatttcgtttgata 173

RESULT 7
US-60-127-461-5765/C
: Sequence 5765, Application US/60127461
: GENERAL INFORMATION:
: APPLICANT: Chapman, Rowan
: APPLICANT: Thornton, Michael
: TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF CANDIDA ALBICANS
: FILE REFERENCE: PM-0007-2 P
: CURRENT APPLICATION NUMBER: US/60/127,461
: NUMBER OF SEQ ID NOS: 8247
: SOFTWARE: PERL Program
: SEQ ID NO 5765
: LENGTH: 256
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: AT808589484
: US-60-127-461-5765

Query Match          63.6%; Score 14; DB 45; Length 256;
Best Local Similarity 63.6%; Pred. No. 1.3e+03;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 tcttgagannnnnnnttgata 22
      |||||
Db 158 tcttgagatttcgtttgata 137

RESULT 8
US-08-812-505-2536
: Sequence 2536, Application US/08812505
: GENERAL INFORMATION:
: APPLICANT: Gooding, Douglas H.
: APPLICANT: Stuve, Laura L.
: APPLICANT: Stuart, Susan G.
: APPLICANT: Ito, Laura Y.
: APPLICANT: Akerblom, Ingrid E.
: APPLICANT: Deleage, Angelo M.
: APPLICANT: Naughton, Rebecca E.
: APPLICANT: Klingler, Tod M.
: APPLICANT: Kington, Michael S.
: APPLICANT: Hillman, Jennifer L.
: TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
: NUMBER OF SEQUENCES: 3896
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
: STREET: 3174 PORTER DRIVE
: CITY: PALO ALTO
```

STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/812,505
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/013,529
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: LUTHER, BARBARA J.
REGISTRATION NUMBER: 33954
REFERENCE/DOCKET NUMBER: PD-0126P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 855-0555
TELEFAX: (415) 852-0195
INFORMATION FOR SEQ ID NO: 2536:
SEQUENCE CHARACTERISTICS:
LENGTH: 259 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
CLONE: 865159
US-08-812-505-2536

Query Match 63.6%; Score 14; DB 12; Length 259;
Best Local Similarity 63.6%; Pred. No. 1.3e+03;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 ttttgannnnnnnnnlttgata 22
|||||
DB 32 ttttgannnnnnnnnlttgata 53

RESULT 9
US-08-978-620-2536
; Sequence 2536, Application US/08978620
; GENERAL INFORMATION:
; APPLICANT: Gooding, Douglas H.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Akerblom, Ingrid E.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Naughton, Rebecca E.
; APPLICANT: Klingler, Tod M.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
; TITLE OF INVENTION: HUMAN BRAIN
; NUMBER OF SEQUENCES: 7797
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/978,620
; FILING DATE: HERBWITH

CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/032,151
FILING DATE: DECEMBER 6, 1996
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/812,505
FILING DATE: MARCH 7, 1997
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: LUTHER, BARBARA J.
REGISTRATION NUMBER: 33954
REFERENCE/DOCKET NUMBER: PD-0126-1 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 852-0195
INFORMATION FOR SEQ ID NO: 2536:
SEQUENCE CHARACTERISTICS:
LENGTH: 259 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
CLONE: 865159
US-08-978-620-2536

Query Match 63.6%; Score 14; DB 13; Length 259;
Best Local Similarity 63.6%; Pred. No. 1.3e+03;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 ttttgannnnnnnnnlttgata 22
|||||
DB 32 ttttgannnnnnnnnlttgata 53

RESULT 10
US-60-013-529-2536
; Sequence 2536, Application US/60013529
; GENERAL INFORMATION:
; APPLICANT: Gooding, Douglas H.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Akerblom, Ingrid E.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Naughton, Rebecca E.
; APPLICANT: Klingler, Tod M.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
; TITLE OF INVENTION: ASTROCYTOMA
; NUMBER OF SEQUENCES: 3896
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/60/013,529
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: LUTHER, BARBARA J.
; REGISTRATION NUMBER: 33954
; REFERENCE/DOCKET NUMBER: PD-0126P
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 855-0555
TELEFAX: (415) 852-0195
INFORMATION FOR SEQ ID NO: 2536:
SEQUENCE CHARACTERISTICS:
LENGTH: 259 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
CLONE: 865159
US-60-013-529-2536

Query Match 63.6%; Score 14; DB 34; Length 259;
Best Local Similarity 63.6%; Pred. No. 1.3e+03;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 tgttgannnnnnnttgata 22
|||||
Db 32 TGTTCGATTTGTCATTGTGATA 53

RESULT 11
US-09-411-077-732
Sequence 732, Application US/09411077
GENERAL INFORMATION:
APPLICANT: Goralski, Thomas J.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM RAT LIVER
FILE REFERENCE: P2-0096 US
CURRENT APPLICATION NUMBER: US/09/411.077
EARLIER FILING DATE: 1999-10-04
EARLIER APPLICATION NUMBER: 60/103.748
EARLIER FILING DATE: October 9, 1998
EARLIER APPLICATION NUMBER: 60/111.900
EARLIER FILING DATE: December 10, 1998
NUMBER OF SEQ ID NOS: 1651
SOFTWARE: PERL Program
SEQ ID NO 732
LENGTH: 261
TYPE: DNA
ORGANISM: Rattus norvegicus
FEATURE: -
OTHER INFORMATION: 701257277H1
US-09-411-077-732

Query Match 63.6%; Score 14; DB 18; Length 261;
Best Local Similarity 63.6%; Pred. No. 1.3e+03;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 tgttgannnnnnnttgata 22
|||||
Db 202 tgttgaaaatcctcttgata 223

RESULT 12
US-09-540-213-52141
Sequence 52141, Application US/09540213
GENERAL INFORMATION:
APPLICANT: Seilhamer, Jeffrey J.
APPLICANT: Deleage, Angelo M.
APPLICANT: Stuart, Susan G.
APPLICANT: Stuve, Laura L.
APPLICANT: Mullaly, Sara J.
APPLICANT: Naughton, Rebecca E.
TITLE OF INVENTION: POLYNUCLEOTIDES OF LIVER AND PANCREAS
FILE REFERENCE: PD-1031 CIP
CURRENT APPLICATION NUMBER: US/09/540.213
CURRENT FILING DATE: 2000-03-31
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 53844

SEQ ID NO 52141
LENGTH: 261
TYPE: DNA
ORGANISM: Rattus norvegicus
FEATURE:
NAME/KEY: misc-feature
OTHER INFORMATION: Incyte ID No: rat00178536
US-09-540-213-52141

Query Match 63.6%; Score 14; DB 21; Length 261;
Best Local Similarity 63.6%; Pred. No. 1.3e+03;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 tgttgannnnnnnttgata 22
|||||
Db 202 tgttgaaaatcctcttgata 223

RESULT 13
US-60-103-748-618
Sequence 618, Application US/60103748
GENERAL INFORMATION:
APPLICANT: Stuve, Laura L.
APPLICANT: Goralski, Thomas J.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM TREATED RAT LIVER
FILE REFERENCE: P2-0096 P
CURRENT APPLICATION NUMBER: US/60/103.748
CURRENT FILING DATE: 1998-10-09
NUMBER OF SEQ ID NOS: 1537
SOFTWARE: PERL Program
SEQ ID NO 618
LENGTH: 261
TYPE: DNA
ORGANISM: Rattus norvegicus
FEATURE: -
OTHER INFORMATION: 701257277H1
US-60-103-748-618

Query Match 63.6%; Score 14; DB 43; Length 261;
Best Local Similarity 63.6%; Pred. No. 1.3e+03;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 tgttgannnnnnnttgata 22
|||||
Db 202 tgttgaaaatcctcttgata 223

RESULT 14
US-09-770-255-141
Sequence 141, Application US/09770255
GENERAL INFORMATION:
APPLICANT: Gorlach, Jorn
APPLICANT: An, Yong-Qiang
APPLICANT: Hamilton, Carol M.
APPLICANT: Price, Jennifer L.
APPLICANT: Raines, Tracy M.
APPLICANT: Yu, Yang
APPLICANT: Rameaka, Joshua G.
APPLICANT: Page, Amy
APPLICANT: Mathew, Abraham V.
APPLICANT: Ledford, Brooke L.
APPLICANT: Woessner, Jeffrey P.
APPLICANT: Haas, William David
APPLICANT: Garcia, Carlos A.
APPLICANT: Krickler, Maja
APPLICANT: Slader, Ted
APPLICANT: Davis, Keith R.
APPLICANT: Allen, Keith
APPLICANT: Hoffman, Neil
APPLICANT: Hurban, Patrick

```

; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; TITLE OF INVENTION: thaliana
; FILE REFERENCE: 2030US (PARA-019PRV)
; CURRENT APPLICATION NUMBER: US/09/770,255
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,504
; PRIOR FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 141
; LENGTH: 317
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(317)
; OTHER INFORMATION: n = A,T,C or G
US-09-770-255-141

```

```

Query Match          63.6%; Score 14; DB 30; Length 317;
Best Local Similarity 63.6%; Pred. No. 1.3e+03;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```

```

OY 1 tgttgagannnnnnnlttgata 22
    |||||  |||||
Db 167 tgttgatgtatggtgttgata 188

```

```

RESULT 15
US-09-705-926-17256/C
; Sequence 17256, Application US/09705926
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: ANNOTATED PLANT GENES
; FILE REFERENCE: 38-21(15481)A
; CURRENT APPLICATION NUMBER: US/09/705,926
; CURRENT FILING DATE: 2000-11-06
; NUMBER OF SEQ ID NOS: 21634
; SEQ ID NO: 17256
; LENGTH: 322
; TYPE: DNA
; ORGANISM: Lycopersicon esculentum
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(322)
; OTHER INFORMATION: unsure at all n locations
US-09-705-926-17256

```

```

Query Match          63.6%; Score 14; DB 28; Length 322;
Best Local Similarity 63.6%; Pred. No. 1.3e+03;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```

```

OY 1 tgttgagannnnnnnlttgata 22
    |||||  |||||
Db 112 TGTTCGAGAGGTTCTTTTGATA 91

```

Search completed: December 15, 2001, 05:42:59
 Job time: 15824 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 05:58:20 ; Search time 474.83 Seconds
(without alignments)
66.629 Million cell updates/sec

Title: US-09-380-826a-6

Perfect score: 1 lgttgannnnnnnttgata 22

Sequence:

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 973175 seqs, 719035916 residues

Total number of hits satisfying chosen parameters: 1946350

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/1/pna/PCR_NEW_COMB.seq:*
2: /cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq:*
3: /cgn2_6/ptodata/1/pna/US07_NEW_COMB.seq:*
4: /cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq:*
5: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:*
6: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq1:*
7: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:*
8: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	14	63.6	427	6	US-09-985-678-90042
C 2	14	63.6	461	6	US-09-985-678-85089
C 3	14	63.6	857	6	US-09-922-279A-3008
C 4	14	63.6	1287	1	PCT-US01-15674A-50
C 5	14	63.6	1501	6	US-09-815-264-12871
C 6	14	63.6	1501	6	US-09-815-264-14471
C 7	14	63.6	1501	6	US-09-815-264-50077
C 8	14	63.6	1501	6	US-09-815-264-50328
C 9	14	63.6	6926	6	US-09-815-264-60943
C 10	14	63.6	7316	6	US-09-815-264-65885
C 11	14	63.6	7673	6	US-09-815-264-74657
C 12	14	63.6	41742	6	US-09-815-264-59220
C 13	14	63.6	94612	6	US-09-815-264-79259
C 14	14	59.1	226	6	US-09-815-242-1444
C 15	14	59.1	245	6	US-09-894-949A-151011
C 16	14	59.1	300	6	US-09-894-949A-1500
C 17	14	59.1	320	6	US-09-815-242-3155
C 18	14	59.1	373	6	US-09-815-242-2111
C 19	14	59.1	378	6	US-09-985-678-231502
C 20	14	59.1	388	6	US-09-388-906A-23742
C 21	14	59.1	425	6	US-09-933-524A-18588
C 22	14	59.1	424	6	US-09-991-423-606
C 23	14	59.1	443	6	US-09-933-524A-110244
C 24	14	59.1	443	6	US-09-815-264-29506
C 25	14	59.1	464	6	US-09-815-242-1944

26	13	59.1	471	6	US-09-933-524A-109421	Sequence 109421,
C 27	13	59.1	501	6	US-09-985-678-220277	Sequence 220277,
C 28	13	59.1	709	6	US-09-815-264-65443	Sequence 65443, A
C 29	13	59.1	845	6	US-09-922-279A-1268	Sequence 1268, Ap
C 30	13	59.1	918	6	US-09-815-242-4840	Sequence 4840, Ap
C 31	13	59.1	1092	6	US-09-922-279A-3103	Sequence 3103, Ap
C 32	13	59.1	1263	6	US-09-815-242-4274	Sequence 4274, Ap
C 33	13	59.1	1266	6	US-09-815-242-8090	Sequence 8090, Ap
C 34	13	59.1	1266	6	US-09-815-242-8747	Sequence 8747, Ap
C 35	13	59.1	1415	6	US-09-898-888A-14137	Sequence 14137, A
C 36	13	59.1	1481	6	US-09-815-264-106283	Sequence 106283,
C 37	13	59.1	1501	6	US-09-815-264-7540	Sequence 7540, Ap
C 38	13	59.1	1501	6	US-09-815-264-25196	Sequence 25196, A
C 39	13	59.1	1537	6	US-09-815-264-106280	Sequence 106280,
C 40	13	59.1	3209	6	US-09-897-516-970	Sequence 970, App
C 41	13	59.1	3498	6	US-09-815-242-6647	Sequence 6647, Ap
C 42	13	59.1	3627	6	US-09-815-264-91108	Sequence 91108, A
C 43	13	59.1	7391	6	US-09-815-264-58271	Sequence 58271, A
C 44	13	59.1	9854	6	US-09-815-264-66587	Sequence 66587, A
C 45	13	59.1	10040	6	US-09-815-264-78579	Sequence 78579, A

ALIGNMENTS

```
RESULT 1
US-09-985-678-90042/C
Sequence 90042, Application US/09985678
GENERAL INFORMATION:
APPLICANT: Cheikh, Nordin
TITLE OF INVENTION: Annotated Plant Genes
FILE REFERENCE: 16517.255/38-21(15097)F
CURRENT FILING DATE: 2001-11-05
PRIOR APPLICATION NUMBER: US 09/304, 517
PRIOR FILING DATE: 1999-05-06
NUMBER OF SEQ ID NOS: 295529
SEQ ID NO 90042
LENGTH: 427
TYPE: DNA
ORGANISM: Zea mays
US-09-985-678-90042

Query Match      63.6% Score 14; DB 6; Length 427;
Best Local Similarity 63.6% Pred. NO. 2.2e+02;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 lgttgannnnnnnttgata 22
|||||
Db 262 TGTTCGATGCACAGTTTGATA 241

RESULT 2
US-09-985-678-85089/C
Sequence 85089, Application US/09985678
GENERAL INFORMATION:
APPLICANT: Cheikh, Nordin
TITLE OF INVENTION: Annotated Plant Genes
FILE REFERENCE: 16517.255/38-21(15097)F
CURRENT FILING DATE: 2001-11-05
PRIOR APPLICATION NUMBER: US 09/304, 517
PRIOR FILING DATE: 1999-05-06
NUMBER OF SEQ ID NOS: 295529
SEQ ID NO 85089
LENGTH: 461
TYPE: DNA
ORGANISM: Zea mays
US-09-985-678-85089
```

Query Match 63.6%; Score 14; DB 6; Length 461;
Best Local Similarity 63.6%; Pred. No. 2.2e+02;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 tgttgagannnnnnnttgata 22
|||||||
DB 300 TGTGGAGAGACAGTTTGATA 279

RESULT 3

US-09-922-279A-3008
; Sequence 3008, Application US/09922279A
; GENERAL INFORMATION:
; APPLICANT: Tang, Yuanhua T.
; APPLICANT: Tillinghast, John
; APPLICANT: Slinku, Ankura
; APPLICANT: Liu, Chenghua
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Contigs Obtained
; FILE REFERENCE: 785
; CURRENT APPLICATION NUMBER: US/09/922,279A
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/491,404
; PRIOR FILING DATE: 2000-01-25
; NUMBER OF SEQ ID NOS: 3796
; SOFTWARE: PL_SP_genes Version 1.0
; SEQ ID NO 3008
; LENGTH: 857
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: (514)...(649)
; OTHER INFORMATION: this location contains the signal peptide sequence.
; OTHER INFORMATION: MGAFTVRLPCSFAMCACLILGPASALCSLCMKMFQFFVILDS, Run with SignalP
; NAME/KEY: misc_feature
; LOCATION: (290)...(213)
; OTHER INFORMATION: similar to g12231098 in the genepept database release 114,
; OTHER INFORMATION: Run with FASTRY 3.3c00, default parameters
US-09-922-279A-3008

Query Match 63.6%; Score 14; DB 6; Length 857;
Best Local Similarity 63.6%; Pred. No. 2.4e+02;

Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 tgttgagannnnnnnttgata 22
|||||||
DB 685 tgttgagaggaagtcgttgata 706

RESULT 4

PCT-US01-15674A-50
; Sequence 50, Application PC/TUS0115674A
; GENERAL INFORMATION:
; APPLICANT: JOAN, KNOUL H
; APPLICANT: ROGAN, PETER K
; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATING
; FILE REFERENCE: 30307
; CURRENT APPLICATION NUMBER: PCT/US01/15674A
; CURRENT FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 613
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 50
; LENGTH: 1287
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: (1)...(1287)
; OTHER INFORMATION: hsmarl

PCT-US01-15674A-50

Query Match 63.6%; Score 14; DB 1; Length 1287;
Best Local Similarity 63.6%; Pred. No. 2.5e+02;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 tgttgagannnnnnnttgata 22
|||||||
DB 36 tgttgagattgcgcttgata 57

RESULT 5

US-09-815-264-12871
; Sequence 12871, Application US/09815264
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Cao, Yongwei
; APPLICANT: Dotson, Stanton B.
; APPLICANT: Koshi, Jeffrey M.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: McIninch, James
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51237)G
; CURRENT APPLICATION NUMBER: US/09/815,264
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 09/620,392
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/702,134
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 109669
; SEQ ID NO 12871
; LENGTH: 1501
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-815-264-12871

Query Match 63.6%; Score 14; DB 6; Length 1501;
Best Local Similarity 63.6%; Pred. No. 2.6e+02;

Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 tgttgagannnnnnnttgata 22
|||||||
DB 671 tgttgagaggaatttgata 692

RESULT 6

US-09-815-264-14471
; Sequence 14471, Application US/09815264
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Cao, Yongwei
; APPLICANT: Dotson, Stanton B.
; APPLICANT: Koshi, Jeffrey M.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: McIninch, James
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51237)G
; CURRENT APPLICATION NUMBER: US/09/815,264
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 09/620,392
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/702,134
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 109669
; SEQ ID NO 14471
; LENGTH: 1501
; TYPE: DNA

ORGANISM: Oryza sativa
US-09-815-264-14471

Query Match 63.6%; Score 14; DB 6; Length 1501;
Best Local Similarity 63.6%; Pred. No. 2.6e+02;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 tgttgannnnnnnttgata 22
||||| |||||||
Db 671 tgttgacaggaatttgata 692

RESULT 7
US-09-815-264-50077/C
Sequence 50077, Application US/09815264

GENERAL INFORMATION:
APPLICANT: Boukharov, Andrey A.
APPLICANT: Cao, Yongwei
APPLICANT: Dotson, Stanton B.
APPLICANT: Koshi, Jeffrey M.
APPLICANT: Kovalic, David K.
APPLICANT: Liu, Jindong
APPLICANT: McIninch, James
APPLICANT: Wu, Wei
TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof
FILE REFERENCE: 38-21(51237)G
CURRENT APPLICATION NUMBER: US/09/815,264
CURRENT FILING DATE: 2001-03-23
PRIOR APPLICATION NUMBER: US 09/620,392
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: US 09/702,134
PRIOR FILING DATE: 2000-10-31
NUMBER OF SEQ ID NOS: 109669
SEQ ID NO 50077
LENGTH: 1501
TYPE: DNA
ORGANISM: Oryza sativa
US-09-815-264-50077

Query Match 63.6%; Score 14; DB 6; Length 1501;
Best Local Similarity 63.6%; Pred. No. 2.6e+02;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 tgttgannnnnnnttgata 22
||||| |||||||
Db 1311 TGTGATATTATCTTGATA 1290

RESULT 8
US-09-815-264-50328
Sequence 50328, Application US/09815264

GENERAL INFORMATION:
APPLICANT: Boukharov, Andrey A.
APPLICANT: Cao, Yongwei
APPLICANT: Dotson, Stanton B.
APPLICANT: Koshi, Jeffrey M.
APPLICANT: Kovalic, David K.
APPLICANT: Liu, Jindong
APPLICANT: McIninch, James
APPLICANT: Wu, Wei
TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof
FILE REFERENCE: 38-21(51237)G
CURRENT APPLICATION NUMBER: US/09/815,264
CURRENT FILING DATE: 2001-03-23
PRIOR APPLICATION NUMBER: US 09/620,392
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: US 09/702,134
PRIOR FILING DATE: 2000-10-31
NUMBER OF SEQ ID NOS: 109669
SEQ ID NO 50328
LENGTH: 1501

TYPE: DNA
ORGANISM: Oryza sativa
US-09-815-264-50328

Query Match 63.6%; Score 14; DB 6; Length 1501;
Best Local Similarity 63.6%; Pred. No. 2.6e+02;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 tgttgannnnnnnttgata 22
||||| |||||||
Db 124 tgttgataatccttgata 145

RESULT 9
US-09-815-264-60943
Sequence 60943, Application US/09815264

GENERAL INFORMATION:
APPLICANT: Boukharov, Andrey A.
APPLICANT: Cao, Yongwei
APPLICANT: Dotson, Stanton B.
APPLICANT: Koshi, Jeffrey M.
APPLICANT: Kovalic, David K.
APPLICANT: Liu, Jindong
APPLICANT: McIninch, James
APPLICANT: Wu, Wei
TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof
FILE REFERENCE: 38-21(51237)G
CURRENT APPLICATION NUMBER: US/09/815,264
CURRENT FILING DATE: 2001-03-23
PRIOR APPLICATION NUMBER: US 09/620,392
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: US 09/702,134
PRIOR FILING DATE: 2000-10-31
NUMBER OF SEQ ID NOS: 109669
SEQ ID NO 60943
LENGTH: 6926
TYPE: DNA
ORGANISM: Oryza sativa
US-09-815-264-60943

Query Match 63.6%; Score 14; DB 6; Length 6926;
Best Local Similarity 63.6%; Pred. No. 3.2e+02;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 tgttgannnnnnnttgata 22
||||| |||||||
Db 4746 tgttgacaggaatttgata 4767

RESULT 10
US-09-815-264-69885/C
Sequence 69885, Application US/09815264

GENERAL INFORMATION:
APPLICANT: Boukharov, Andrey A.
APPLICANT: Cao, Yongwei
APPLICANT: Dotson, Stanton B.
APPLICANT: Koshi, Jeffrey M.
APPLICANT: Kovalic, David K.
APPLICANT: Liu, Jindong
APPLICANT: McIninch, James
APPLICANT: Wu, Wei
TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof
FILE REFERENCE: 38-21(51237)G
CURRENT APPLICATION NUMBER: US/09/815,264
CURRENT FILING DATE: 2001-03-23
PRIOR APPLICATION NUMBER: US 09/620,392
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: US 09/702,134
PRIOR FILING DATE: 2000-10-31
NUMBER OF SEQ ID NOS: 109669
SEQ ID NO 69885

LENGTH: 7316
TYPE: DNA
ORGANISM: Oryza sativa
US-09-815-264-69885

Query Match 63.6%; Score 14; DB 6; Length 7316;
Best Local Similarity 63.6%; Pred. No. 3.2e+02;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 tgttgannnnnnnttgata 22
|||||
Db 3151 TGTGGATTGTTGATTGATA 3130

RESULT 11
US-09-815-264-74657/C
Sequence 74657, Application US/09815264
GENERAL INFORMATION:
APPLICANT: Boukharov, Andrey A.
APPLICANT: Cao, Yongwei
APPLICANT: Dotson, Stanton B.
APPLICANT: Koshi, Jeffrey M.
APPLICANT: Kovalic, David K.
APPLICANT: Liu, Jingdong
APPLICANT: McIninch, James
APPLICANT: Wu, Wei
TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof
FILE REFERENCE: 38-21(51237)G
CURRENT APPLICATION NUMBER: US/09/815,264
CURRENT FILING DATE: 2001-03-23
PRIOR APPLICATION NUMBER: US 09/620,392
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: US 09/702,134
PRIOR FILING DATE: 2000-10-31
NUMBER OF SEQ ID NOS: 109669
SEQ ID NO 74657
LENGTH: 7673
TYPE: DNA
ORGANISM: Oryza sativa
US-09-815-264-74657

Query Match 63.6%; Score 14; DB 6; Length 7673;
Best Local Similarity 63.6%; Pred. No. 3.2e+02;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 tgttgannnnnnnttgata 22
|||||
Db 3586 TGTGGACATATATTGATA 3565

RESULT 12
US-09-815-264-59220
Sequence 59220, Application US/09815264
GENERAL INFORMATION:
APPLICANT: Boukharov, Andrey A.
APPLICANT: Cao, Yongwei
APPLICANT: Dotson, Stanton B.
APPLICANT: Koshi, Jeffrey M.
APPLICANT: Kovalic, David K.
APPLICANT: Liu, Jingdong
APPLICANT: McIninch, James
APPLICANT: Wu, Wei
TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof
FILE REFERENCE: 38-21(51237)G
CURRENT APPLICATION NUMBER: US/09/815,264
CURRENT FILING DATE: 2001-03-23
PRIOR APPLICATION NUMBER: US 09/620,392
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: US 09/702,134
PRIOR FILING DATE: 2000-10-31
NUMBER OF SEQ ID NOS: 109669

SEQ ID NO 59220
LENGTH: 41742
TYPE: DNA
ORGANISM: Oryza sativa
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(41742)
OTHER INFORMATION: unsure at all n locations
US-09-815-264-59220

Query Match 63.6%; Score 14; DB 6; Length 41742;
Best Local Similarity 63.6%; Pred. No. 4.1e+02;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 tgttgannnnnnnttgata 22
|||||
Db 2402 tgttgaaacagaatttgata 2423

RESULT 13
US-09-815-264-79259
Sequence 79259, Application US/09815264
GENERAL INFORMATION:
APPLICANT: Boukharov, Andrey A.
APPLICANT: Cao, Yongwei
APPLICANT: Dotson, Stanton B.
APPLICANT: Koshi, Jeffrey M.
APPLICANT: Kovalic, David K.
APPLICANT: Liu, Jingdong
APPLICANT: McIninch, James
APPLICANT: Wu, Wei
TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof
FILE REFERENCE: 38-21(51237)G
CURRENT APPLICATION NUMBER: US/09/815,264
CURRENT FILING DATE: 2001-03-23
PRIOR APPLICATION NUMBER: US 09/620,392
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: US 09/702,134
PRIOR FILING DATE: 2000-10-31
NUMBER OF SEQ ID NOS: 109669
SEQ ID NO 79259
LENGTH: 94612
TYPE: DNA
ORGANISM: Oryza sativa
US-09-815-264-79259

Query Match 63.6%; Score 14; DB 6; Length 94612;
Best Local Similarity 63.6%; Pred. No. 4.5e+02;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 tgttgannnnnnnttgata 22
|||||
Db 14796 tgttgatattatccttgata 14817

RESULT 14
US-09-815-242-1444
Sequence 1444, Application US/09815242
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242


```

; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1444
; LENGTH: 226
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-09-815-242-1444

```

```

Query Match          59.1%; Score 13; DB 6; Length 226;
Best Local Similarity 61.9%; Pred. No. 6.9e+02;
Matches 13; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 2 gtggannnnnnnnttgata 22
    |||||
Db 122 gtggatgtagcttcttgata 142

```

```

RESULT 15
US-09-985-678-151011/C
; Sequence 151011, Application US/09985678
; GENERAL INFORMATION:
; APPLICANT: Cheikh, Nordine
; TITLE OF INVENTION: Annotated Plant Genes
; FILE REFERENCE: 16517.255/38-21(15087)P
; CURRENT APPLICATION NUMBER: US/09/985,678
; CURRENT FILING DATE: 2001-11-05
; PRIOR APPLICATION NUMBER: US 09/304,517
; PRIOR FILING DATE: 1999-05-06
; NUMBER OF SEQ ID NOS: 295529
; SEQ ID NO 151011
; LENGTH: 245
; TYPE: DNA
; ORGANISM: Glycine max
US-09-985-678-151011

```

```

Query Match          59.1%; Score 13; DB 6; Length 245;
Best Local Similarity 61.9%; Pred. No. 6.9e+02;
Matches 13; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 2 gtggannnnnnnnttgata 22
    |||||
Db 165 GTTGCATTGTCATCTTGATA 145

```

Search completed: December 15, 2001, 05:58:24
Job time: 12267 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 05:42:59 ; Search time 7962.77 Seconds
(without alignments)
48,423 Million cell updates/sec

Title: US-09-380-826a-7

Perfect score: 1 ttttgatcacacaagatttgata 22

Sequence: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 17159718 seqs, 8763200856 residues

Total number of hits satisfying chosen parameters: 34319436

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_Main:*

```
1: /cgn2_6/ptodata/2/pna/US0611_COMB.seq:*
2: /cgn2_6/ptodata/2/pna/US06_COMB.seq:*
3: /cgn2_6/ptodata/2/pna/US07_COMB.seq:*
4: /cgn2_6/ptodata/2/pna/US080_COMB.seq:*
5: /cgn2_6/ptodata/2/pna/US081_COMB.seq:*
6: /cgn2_6/ptodata/2/pna/US082_COMB.seq:*
7: /cgn2_6/ptodata/2/pna/US083_COMB.seq:*
8: /cgn2_6/ptodata/2/pna/US084_COMB.seq:*
9: /cgn2_6/ptodata/2/pna/US085_COMB.seq:*
10: /cgn2_6/ptodata/2/pna/US086_COMB.seq:*
11: /cgn2_6/ptodata/2/pna/US087_COMB.seq:*
12: /cgn2_6/ptodata/2/pna/US088_COMB.seq:*
13: /cgn2_6/ptodata/2/pna/US089_COMB.seq:*
14: /cgn2_6/ptodata/2/pna/US090_COMB.seq:*
15: /cgn2_6/ptodata/2/pna/US091_COMB.seq:*
16: /cgn2_6/ptodata/2/pna/US092_COMB.seq:*
17: /cgn2_6/ptodata/2/pna/US093_COMB.seq:*
18: /cgn2_6/ptodata/2/pna/US094_COMB.seq:*
19: /cgn2_6/ptodata/2/pna/US095_COMB.seq:*
20: /cgn2_6/ptodata/2/pna/US095C_COMB.seq:*
21: /cgn2_6/ptodata/2/pna/US095D_COMB.seq:*
22: /cgn2_6/ptodata/2/pna/US096A_COMB.seq:*
23: /cgn2_6/ptodata/2/pna/US096B_COMB.seq:*
24: /cgn2_6/ptodata/2/pna/US096C_COMB.seq:*
25: /cgn2_6/ptodata/2/pna/US096D_COMB.seq:*
26: /cgn2_6/ptodata/2/pna/US096E_COMB.seq:*
27: /cgn2_6/ptodata/2/pna/US096F_COMB.seq:*
28: /cgn2_6/ptodata/2/pna/US097A_COMB.seq:*
29: /cgn2_6/ptodata/2/pna/US097B_COMB.seq:*
30: /cgn2_6/ptodata/2/pna/US097C_COMB.seq:*
31: /cgn2_6/ptodata/2/pna/US098_COMB.seq:*
32: /cgn2_6/ptodata/2/pna/US099_COMB.seq:*
33: /cgn2_6/ptodata/2/pna/US6000_COMB.seq:*
34: /cgn2_6/ptodata/2/pna/US6001_COMB.seq:*
35: /cgn2_6/ptodata/2/pna/US6002_COMB.seq:*
36: /cgn2_6/ptodata/2/pna/US6003_COMB.seq:*
37: /cgn2_6/ptodata/2/pna/US6004_COMB.seq:*
38: /cgn2_6/ptodata/2/pna/US6005_COMB.seq:*
39: /cgn2_6/ptodata/2/pna/US6006_COMB.seq:*
40: /cgn2_6/ptodata/2/pna/US6007_COMB.seq:*
41: /cgn2_6/ptodata/2/pna/US6008_COMB.seq:*
42: /cgn2_6/ptodata/2/pna/US6009_COMB.seq:*
43: /cgn2_6/ptodata/2/pna/US6010_COMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
44:	22	100.0	22	17	US-09-380-826a-7	Sequence 2, Appli
45:	22	100.0	22	17	US-09-380-826a-7	Sequence 1, Appli
46:	22	100.0	1477	17	US-09-380-826a-1	Sequence 11944, A
47:	18.8	85.5	420	46	US-60-132-861-11944	Sequence 7425, Ap
48:	18.8	85.5	530	25	US-09-649-165a-7425	Sequence 175385,
49:	18.8	85.5	581	25	US-09-654-617-175385	Sequence 20908, A
50:	18.8	85.5	581	27	US-09-684-016-117747	Sequence 49885, A
51:	18.8	85.5	581	25	US-09-684-016-117747	Sequence 117747,
52:	18.8	85.5	581	25	US-09-684-016-117747	Sequence 117747,
53:	18.8	85.5	581	25	US-09-684-016-117747	Sequence 117747,
54:	18.8	85.5	581	25	US-09-684-016-117747	Sequence 117747,
55:	18.8	85.5	581	25	US-09-684-016-117747	Sequence 117747,
56:	18.8	85.5	581	25	US-09-684-016-117747	Sequence 117747,
57:	18.8	85.5	581	25	US-09-684-016-117747	Sequence 117747,
58:	18.8	85.5	581	25	US-09-684-016-117747	Sequence 117747,
59:	18.8	85.5	581	25	US-09-684-016-117747	Sequence 117747,
60:	18.8	85.5	581	25	US-09-684-016-117747	Sequence 117747,
61:	18.8	85.5	581	25	US-09-684-016-117747	Sequence 117747,
62:	18.8	85.5	581	25	US-09-684-016-117747	Sequence 117747,
63:	18.8	85.5	581	25	US-09-684-016-117747	Sequence 117747,
64:	18.8	85.5	581	25	US-09-684-016-117747	Sequence 117747,
65:	18.8	85.5	581	25	US-09-684-016-117747	Sequence 117747,

42	17.2	78.2	249	24	US-09-638-258-520	Sequence 520, App
43	17.2	78.2	249	32	US-09-924-035-520	Sequence 520, App
44	17.2	78.2	278	25	US-09-654-617-370318	Sequence 370318,
45	17.2	78.2	278	27	US-09-684-016-370318	Sequence 370318,

ALIGNMENTS

```
RESULT 1
US-09-380-826a-7
; Sequence 2, Application US/09380826A
; GENERAL INFORMATION:
; APPLICANT: Chappel, Rod
; TITLE OF INVENTION: LEPTOSPIRA PATHOGENS
; FILE REFERENCE: DAVIE79.001APC
; CURRENT APPLICATION NUMBER: US/09/380,826A
; PRIOR FILING DATE: 1999-11-22
; PRIOR APPLICATION NUMBER: PCT/AU98/00145
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: AU P05494/97
; PRIOR FILING DATE: 1997-03-07
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Primer for L. fainel.
US-09-380-826a-2
```

```
Query Match          100.0%; Score 22; DB 17; Length 22;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 ttttgatcacagaatttgata 22
    |||
Db 1 ttttgatcacagaatttgata 22
```

```
RESULT 2
US-09-380-826a-7
; Sequence 7, Application US/09380826A
; GENERAL INFORMATION:
; APPLICANT: Chappel, Rod
; TITLE OF INVENTION: LEPTOSPIRA PATHOGENS
; FILE REFERENCE: DAVIE79.001APC
; CURRENT APPLICATION NUMBER: US/09/380,826A
; PRIOR FILING DATE: 1999-11-22
; PRIOR APPLICATION NUMBER: PCT/AU98/00145
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: AU P05494/97
; PRIOR FILING DATE: 1997-03-07
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Primer for L. fainel.
US-09-380-826a-7
```

```
Query Match          100.0%; Score 22; DB 17; Length 22;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 ttttgatcacagaatttgata 22
    |||
Db 1 ttttgatcacagaatttgata 22
```

```
RESULT 3
US-09-380-826a-1
; Sequence 1, Application US/09380826A
; GENERAL INFORMATION:
; APPLICANT: Chappel, Rod
; TITLE OF INVENTION: LEPTOSPIRA PATHOGENS
; FILE REFERENCE: DAVIE79.001APC
; CURRENT APPLICATION NUMBER: US/09/380,826A
; PRIOR FILING DATE: 1999-11-22
; PRIOR APPLICATION NUMBER: PCT/AU98/00145
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: AU P05494/97
; PRIOR FILING DATE: 1997-03-07
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1477
; TYPE: DNA
; ORGANISM: Leptospira fainel
US-09-380-826a-1
```

```
Query Match          100.0%; Score 22; DB 17; Length 1477;
Best Local Similarity 100.0%; Pred. No. 3,3;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 ttttgatcacagaatttgata 22
    |||
Db 154 ttttgatcacagaatttgata 175
```

```
RESULT 4
US-60-132-861-11944
; Sequence 11944, Application US/60132861
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Timberlake, William E.
; TITLE OF INVENTION: ASPERGILLUS NIDULANS GENOME SEQUENCE AND USES THEREOF
; FILE REFERENCE: 38-21(15485)C
; CURRENT APPLICATION NUMBER: US/60/132,861
; CURRENT FILING DATE: 1999-05-07
; NUMBER OF SEQ ID NOS: 13800
; SEQ ID NO 11944
; LENGTH: 420
; TYPE: DNA
; ORGANISM: Aspergillus nidulans
US-60-132-861-11944
```

```
Query Match          85.5%; Score 18.8; DB 46; Length 420;
Best Local Similarity 90.9%; Pred. No. 1,1e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 1 ttttgatcacagaatttgata 22
    |||
Db 234 ttttgatcacagaatttgata 255
```

```
RESULT 5
US-09-649-165a-7425/c
; Sequence 7425, Application US/09649165A
; GENERAL INFORMATION:
; APPLICANT: Guegler, Karl J.
; APPLICANT: Doyle, Martin
; APPLICANT: Momiyama, Monika
; APPLICANT: Wang, Xinhao
; TITLE OF INVENTION: DETECTING EXPRESSION OF ARABIDOPSIS THALIANA GENES
; FILE REFERENCE: PA-0019 US
; CURRENT APPLICATION NUMBER: US/09/649,165A
; CURRENT FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 60/150,517
```

PRIOR FILING DATE: 1999-08-24
NUMBER OF SEQ ID NOS: 7753
SOFTWARE: PERL Program
SEQ ID NO 7425
LENGTH: 530
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: Incycle ID No: 701676626
NAME/KEY: unsure
LOCATION: 6
OTHER INFORMATION: a, t, c, g, dr other
US-09-649-165A-7425

Query Match 85.5%; Score 18.8; DB 25; Length 530;
Best Local Similarity 90.9%; Pred. No. 1.1e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 tgttgatcacagaatttgata 22
||||| |||||||||
Db 254 TGTGTGTTCAACAAGATTTCACA 233

RESULT 6
US-09-654-617-175385
Sequence 175385, Application US/09654617
GENERAL INFORMATION:
APPLICANT: Kovallc, David K.
APPLICANT: Liu, Jindong
TITLE OF INVENTION: Annotated Plant Genes
FILE REFERENCE: 38-21(15097)D
CURRENT APPLICATION NUMBER: US/09/654,617
CURRENT FILING DATE: 2000-09-05
NUMBER OF SEQ ID NOS: 463173
SEQ ID NO 175385
LENGTH: 581
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-654-617-175385

Query Match 85.5%; Score 18.8; DB 25; Length 581;
Best Local Similarity 90.9%; Pred. No. 1.1e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 tgttgatcacagaatttgata 22
||||| |||||||||
Db 420 tgttgatcacagaatttgata 441

RESULT 7
US-09-684-016-175385
Sequence 175385, Application US/09684016
GENERAL INFORMATION:
APPLICANT: Kovallc, David K.
APPLICANT: Liu, Jindong
TITLE OF INVENTION: Annotated Plant Genes
FILE REFERENCE: 38-21(15097)D
CURRENT APPLICATION NUMBER: US/09/684,016
CURRENT FILING DATE: 2000-10-10
PRIOR APPLICATION NUMBER: US 09/654,617
PRIOR FILING DATE: 2000-09-05
NUMBER OF SEQ ID NOS: 463173
SEQ ID NO 175385
LENGTH: 581
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-684-016-175385

Query Match 85.5%; Score 18.8; DB 27; Length 581;

Best Local Similarity 90.9%; Pred. No. 1.1e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 tgttgatcacagaatttgata 22
||||| |||||||||
Db 420 tgttgatcacagaatttgata 441

RESULT 8
US-09-565-309A-20908
Sequence 20908, Application US/09565309A
GENERAL INFORMATION:
APPLICANT: ALEXANDROV, Nikolai
APPLICANT: BROVER, Vyacheslav
TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
FILE REFERENCE: 2750-0853p
CURRENT APPLICATION NUMBER: US/09/565,309A
CURRENT FILING DATE: 2000-05-05
NUMBER OF SEQ ID NOS: 68449
SEQ ID NO 20908
LENGTH: 649
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1)..(649)
OTHER INFORMATION: any n = a, g, c, t, unknown, or other
NAME/KEY: misc.feature
LOCATION: (1)..(649)
OTHER INFORMATION: 27496:110925 (Clone Number:Unique Sequence Identifier)
US-09-565-309A-20908

Query Match 85.5%; Score 18.8; DB 22; Length 649;
Best Local Similarity 90.9%; Pred. No. 1.1e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 tgttgatcacagaatttgata 22
||||| |||||||||
Db 265 tgttgatcacagaatttgata 286

RESULT 9
US-09-565-309A-57118
Sequence 57118, Application US/09565309A
GENERAL INFORMATION:
APPLICANT: ALEXANDROV, Nikolai
APPLICANT: BROVER, Vyacheslav
TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
FILE REFERENCE: 2750-0853p
CURRENT APPLICATION NUMBER: US/09/565,309A
CURRENT FILING DATE: 2000-05-05
NUMBER OF SEQ ID NOS: 68449
SEQ ID NO 57118
LENGTH: 1069
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1)..(1069)
OTHER INFORMATION: any n = a, g, c, t, unknown, or other
NAME/KEY: misc.feature
LOCATION: (1)..(1069)
OTHER INFORMATION: 27496 : OVERLAP (Clone Number : OVERLAP)
US-09-565-309A-57118

Query Match 85.5%; Score 18.8; DB 22; Length 1069;
Best Local Similarity 90.9%; Pred. No. 1.2e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ttttgatcacaaagtattgata 22
||||| ||||||||| |||
Db 690 ttttgatcacaaagtattgaca 711

RESULT 10

US-09-565-309A-49885
; Sequence 49885, Application US/09565309A
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai
; APPLICANT: BROVER, Vyacheslav
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-0853P
; CURRENT APPLICATION NUMBER: US/09/565,309A
; CURRENT FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 68449
; SEQ ID NO 49885
; LENGTH: 1072
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1072)
; OTHER INFORMATION: any n = a, g, c, t, unknown, or other
; NAME/KEY: misc_feature
; LOCATION: (1)..(1072)
; OTHER INFORMATION: 27496 : 5TAG CONSENSUS (Clone Number:5tag_consensus)
US-09-565-309A-49885

Query Match 85.5%; Score 18.8; DB 22; Length 1072;
Best Local Similarity 90.9%; Pred. No. 1.2e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ttttgatcacaaagtattgata 22
||||| ||||||||| |||
Db 691 ttttgatcacaaagtattgaca 712

RESULT 11

US-09-654-617-117747
; Sequence 117747, Application US/09654617
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; TITLE OF INVENTION: Annotated Plant Genes
; FILE REFERENCE: 38-21(15097)D
; CURRENT APPLICATION NUMBER: US/09/654,617
; CURRENT FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 463173
; SEQ ID NO 117747
; LENGTH: 1128
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; OTHER INFORMATION: unsure at all n locations
US-09-654-617-117747

Query Match 85.5%; Score 18.8; DB 25; Length 1128;
Best Local Similarity 90.9%; Pred. No. 1.2e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ttttgatcacaaagtattgata 22
||||| ||||||||| |||
Db 719 ttttgatcacaaagtattgaca 740

RESULT 12

US-09-684-016-117747
; Sequence 117747, Application US/09684016
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.

; APPLICANT: Liu, Jingdong
; TITLE OF INVENTION: Annotated Plant Genes
; FILE REFERENCE: 38-21(15097)D
; CURRENT APPLICATION NUMBER: US/09/684,016
; CURRENT FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: US 09/654,617
; PRIOR FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 463173
; SEQ ID NO 117747
; LENGTH: 1128
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1128)
; OTHER INFORMATION: unsure at all n locations
US-09-684-016-117747

Query Match 85.5%; Score 18.8; DB 27; Length 1128;
Best Local Similarity 90.9%; Pred. No. 1.2e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ttttgatcacaaagtattgata 22
||||| ||||||||| |||
Db 719 ttttgatcacaaagtattgaca 740

RESULT 13

US-60-138-103-11564/C
; Sequence 11564, Application US/60138103
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Timberlake, William E.
; TITLE OF INVENTION: ASPERGILLUS NIDULANS GENOME SEQUENCE AND USES THEREOF
; FILE REFERENCE: 38-10(15485)D
; CURRENT APPLICATION NUMBER: US/60/138,103
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 28006
; SEQ ID NO 11564
; LENGTH: 1429
; TYPE: DNA
; ORGANISM: Aspergillus nidulans
US-60-138-103-11564

Query Match 85.5%; Score 18.8; DB 46; Length 1429;
Best Local Similarity 90.9%; Pred. No. 1.2e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ttttgatcacaaagtattgata 22
||||| ||||||||| |||
Db 1206 TGTGTGATGACAGATTAGATA 1185

RESULT 14

US-09-404-520-5289
; Sequence 5289, Application US/09404520
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Ghodssi, Azita
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: McIninch, James
; APPLICANT: Timberlake, William E.
; APPLICANT: Yu, Jaehyuk
; TITLE OF INVENTION: EMERICELLA NIDULANS GENOME SEQUENCE AND USES THEREOF
; FILE REFERENCE: 38-10(15498)A
; CURRENT APPLICATION NUMBER: US/09/404,520
; CURRENT FILING DATE: 1999-09-23
; NUMBER OF SEQ ID NOS: 44345
; SEQ ID NO 5289
; LENGTH: 5874
; TYPE: DNA

: ORGANISM: Aspergillus nidulans
US-09-404-520-5289

Query Match 85.5%; Score 18.8; DB 18; Length 5874;
Best Local Similarity 90.9%; Pred. No. 1.4e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 tgttgatcacacaagatttgata 22
||||| ||||| ||||| |||||
Db 248 tgttgatcacacaagatttgata 269

RESULT 15

US-09-534-859-352
: Sequence 352, Application US/09534859
: GENERAL INFORMATION:
: APPLICANT: Bush, David F.
: APPLICANT: Last, Robert L.
: APPLICANT: Levin, Irena M.
: APPLICANT: Norris, Susan R.
: APPLICANT: Parnell, Laurence D.
: APPLICANT: Rounsley, Steven D.
: APPLICANT: Wiegand, Roger C.
: TITLE OF INVENTION: PLANT POLYMORPHIC MARKERS AND USES THEREOF
: FILE REFERENCE: 38-10(15493)B
: CURRENT APPLICATION NUMBER: US/09/534,859
: CURRENT FILING DATE: 2000-03-29
: NUMBER OF SEQ ID NOS: 1127
: SEQ ID NO 352
: LENGTH: 92624
: TYPE: DNA
: ORGANISM: Arabidopsis thaliana
US-09-534-859-352

Query Match 85.5%; Score 18.8; DB 20; Length 92624;
Best Local Similarity 90.9%; Pred. No. 2e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 tgttgatcacacaagatttgata 22
||||| ||||| ||||| |||||
Db 35807 tgttgatcacacaagatttgata 35828

Search completed: December 15, 2001, 05:43:01
Job time: 15826 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 05:58:24 ; Search time 474.83 seconds
(without alignments)
66.629 Million cell updates/sec

Title: US-09-380-826a-7

Perfect score: 1 tlggtgacacaaagattgata 22

Sequence: IDENTITY_NUC

Scoring table: Gapop 10.0, Capext 1.0

Searched: 973175 seqs, 719035916 residues

Total number of hits satisfying chosen parameters: 1946350

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_NA_New:*
1: /cgn2_6/ptodata/1/pna/PCr_NEW_COMB.seq:*
2: /cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq:*
3: /cgn2_6/ptodata/1/pna/US07_NEW_COMB.seq:*
4: /cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq:*
5: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:*
6: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq1:*
7: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:*
8: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	18.8	85.5	1069	US-09-620-394B-4668	Sequence 4668, Ap
2	17.2	78.2	249	US-09-924-035A-520	Sequence 520, App
3	17.2	78.2	622	US-09-985-678-211824	Sequence 211824, Ap
4	17.2	78.2	2038	US-09-923-279A-2773	Sequence 2773, Ap
5	17.2	78.2	46891	US-09-815-264-82805	Sequence 82805, A
6	16.8	76.4	267	US-09-982-402-1036	Sequence 1036, Ap
7	16.8	76.4	551	US-09-849-526A-27340	Sequence 27340, A
8	16.8	76.4	849	US-09-815-264-26466	Sequence 26466, A
9	16.8	76.4	1046	US-09-815-264-63319	Sequence 63319, A
10	16.8	76.4	1245	US-09-815-264-91672	Sequence 91672, A
11	16.8	76.4	1501	US-09-815-264-23140	Sequence 23140, A
12	16.8	76.4	1501	US-09-815-264-46882	Sequence 46882, A
13	16.8	76.4	1501	US-09-815-264-53250	Sequence 53250, A
14	16.8	76.4	6940	US-09-815-264-76889	Sequence 76889, A
15	16.8	76.4	12081	US-09-815-264-81045	Sequence 81045, A
16	16.8	76.4	12198	US-09-815-264-76653	Sequence 76653, A
17	16.8	76.4	41269	US-09-815-264-59808	Sequence 59808, A
18	16.8	76.4	50929	US-09-815-264-61033	Sequence 61033, A
19	16.8	76.4	67216	US-09-815-264-71944	Sequence 71944, A
20	16.4	74.5	360	US-09-388-906A-17536	Sequence 17536, A
21	16.4	74.5	4861	US-09-815-264-109466	Sequence 109466, A
22	16.2	73.6	278	US-09-985-678-279860	Sequence 279860, A
23	16.2	73.6	376	US-09-912-293-18469	Sequence 18469, A
24	16.2	73.6	377	US-09-845-487A-238	Sequence 238, App
25	16.2	73.6	419	US-09-933-524A-2094	Sequence 2094, Ap

26	16.2	73.6	461	6	US-09-849-526A-12135	Sequence 12135, A
27	16.2	73.6	845	6	US-09-923-279A-1268	Sequence 1268, Ap
c 28	16.2	73.6	1251	6	US-09-898-888A-9009	Sequence 9009, Ap
c 29	16.2	73.6	1457	6	US-09-815-264-60838	Sequence 60838, A
c 30	16.2	73.6	1501	6	US-09-815-264-39775	Sequence 39775, A
c 31	16.2	73.6	1501	6	US-09-815-264-39776	Sequence 39776, A
c 32	16.2	73.6	1501	6	US-09-815-264-43531	Sequence 43531, A
c 33	16.2	73.6	3603	6	US-09-815-264-40803	Sequence 40803, A
c 34	16.2	73.6	6237	6	US-09-815-264-72243	Sequence 72243, A
c 35	16.2	73.6	7856	6	US-09-815-264-86401	Sequence 86401, A
c 36	16.2	73.6	9723	6	US-09-815-264-75693	Sequence 75693, A
c 37	16.2	73.6	10839	6	US-09-815-264-78946	Sequence 78946, A
c 38	16.2	73.6	26601	6	US-09-815-264-74669	Sequence 74669, A
c 39	16.2	73.6	33005	6	US-09-815-264-73015	Sequence 73015, A
c 40	16.2	73.6	36227	6	US-09-815-264-68619	Sequence 68619, A
c 41	16.2	72.7	610	6	US-09-985-678-213635	Sequence 213635, A
c 42	16.2	72.7	1381	6	US-09-815-264-108378	Sequence 108378, A
c 43	16.2	72.7	37658	6	US-09-815-264-74259	Sequence 74259, A
c 44	16.2	72.7	39365	6	US-09-815-264-79204	Sequence 79204, A
c 45	15.8	71.8	270	6	US-09-985-678-42368	Sequence 42368, A

ALIGNMENTS

```
RESULT 1
US-09-620-394B-4668
; Sequence 4668, Application US/09620394B
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptide
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 2750-1067P
; CURRENT APPLICATION NUMBER: US/09/620.394B
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 9131
; SEQ ID NO 4668
; LENGTH: 1069
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: 1..1069
; OTHER INFORMATION: any n = a, g, c, t, unknown, or other
; NAME/KEY: misc-feature
; LOCATION: 1..1069
; OTHER INFORMATION: Ceres Seq. ID 1393173
; US-09-620-394B-4668

Query Match 85.5%; Score 18.8; DB 6; Length 1069;
Best Local Similarity 90.9%; Pred. No. 16;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 tlggtgacacaaagattgata 22
Db 690 tlggtgacacaaagattgata 711

RESULT 2
US-09-924-035A-520
; Sequence 520, Application US/09924035A
; GENERAL INFORMATION:
; APPLICANT: Grlach, Jrm
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; TITLE OF INVENTION: thaliana
; FILE REFERENCE: 2011US
; CURRENT APPLICATION NUMBER: US/09/924.035A
; CURRENT FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: US 60/148,784
; PRIOR FILING DATE: 1999-08-13
; NUMBER OF SEQ ID NOS: 900
```

```

; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 520
; LENGTH: 249
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(249)
; OTHER INFORMATION: n = A,T,C or G
US-09-924-035A-520

```

```

Query Match      78.2%; Score 17.2; DB 5; Length 249;
Best Local Similarity 86.4%; Pred. No. 81;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 1 tgttgatcacagaattgata 22
    ||||| ||||| ||||| ||
DB 54 tgttgatcacagaattgata 75

```

```

RESULT 3
US-09-985-678-211824
; Sequence 211824, Application US/09985678
; GENERAL INFORMATION:
; APPLICANT: Cheikh, Norline
; TITLE OF INVENTION: Annotated Plant Genes
; FILE REFERENCE: 16517.255/38-21(15097)F
; CURRENT APPLICATION NUMBER: US/09/985,678
; CURRENT FILING DATE: 2001-11-05
; PRIOR APPLICATION NUMBER: US 09/304,517
; PRIOR FILING DATE: 1999-05-06
; NUMBER OF SEQ ID NOS: 295529
; SEQ ID NO 211824
; LENGTH: 622
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)...(622)
; OTHER INFORMATION: unsure at all n locations
US-09-985-678-211824

```

```

Query Match      78.2%; Score 17.2; DB 6; Length 622;
Best Local Similarity 86.4%; Pred. No. 87;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 1 tgttgatcacagaattgata 22
    ||||| ||||| ||||| ||
DB 520 tgttgatcacagaattgata 541

```

```

RESULT 4
US-09-922-279A-2773
; Sequence 2773, Application US/09922279A
; GENERAL INFORMATION:
; APPLICANT: Tang, Yuanhua T.
; APPLICANT: Tillinghast, John
; APPLICANT: Slinku, Ankura
; APPLICANT: Liu, Chenghua
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Contigs Obtained
; TITLE OF INVENTION: From Various Libraries
; FILE REFERENCE: 785
; CURRENT APPLICATION NUMBER: US/09/922,279A
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/491,404
; PRIOR FILING DATE: 2000-01-25
; NUMBER OF SEQ ID NOS: 3796
; SOFTWARE: PL_SP_genes Version 1.0
; SEQ ID NO 2773

```

```

; LENGTH: 2038
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: (1091)...(956)
; OTHER INFORMATION: this location contains the signal peptide sequence
; OTHER INFORMATION: MLYLLFPGVSYLRSLFGRPIGPISTDFLLFLFSNLDSPLS, Run with Signal
; NAME/KEY: misc_feature
; LOCATION: (605)...(1902)
; OTHER INFORMATION: similar to g15931821 in the genepept database release 114,
; OTHER INFORMATION: Run with FASTX 3.3c00, default parameters
US-09-922-279A-2773

```

```

Query Match      78.2%; Score 17.2; DB 6; Length 2038;
Best Local Similarity 86.4%; Pred. No. 96;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 1 tgttgatcacagaattgata 22
    ||||| ||||| ||||| ||
DB 1929 tgttgatcacagaattgata 1950

```

```

RESULT 5
US-09-815-264-82805
; Sequence 82805, Application US/09815264
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Cao, Yongwei
; APPLICANT: Dotson, Stanton B.
; APPLICANT: Koshi, Jeffrey M.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: McIninch, James
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51237)G
; CURRENT APPLICATION NUMBER: US/09/815,264
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 09/620,392
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/702,134
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 109669
; SEQ ID NO 82805
; LENGTH: 46891
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-815-264-82805

```

```

Query Match      78.2%; Score 17.2; DB 6; Length 46891;
Best Local Similarity 86.4%; Pred. No. 1,3e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 1 tgttgatcacagaattgata 22
    ||||| ||||| ||||| ||
DB 45218 tgttgatcacagaattgata 45239

```

```

RESULT 6
US-09-982-402-1036
; Sequence 1036, Application US/09982402
; GENERAL INFORMATION:
; APPLICANT: [list inventors here]
; TITLE OF INVENTION: [list title here]
; FILE REFERENCE: [list docket ID here]
; CURRENT APPLICATION NUMBER: US/09/982,402
; CURRENT FILING DATE: 2001-10-17
; NUMBER OF SEQ ID NOS: 3553
; SOFTWARE: PERL Program
; SEQ ID NO 1036

```

```

: LENGTH: 267
: TYPE: DNA
: ORGANISM: Zea mays
: FEATURE:
: NAME/KEY: misc_feature
: OTHER INFORMATION: Incyte ID No: 700196524H1
US-09-982-402-1036

```

```

Query Match          76.4% Score 16.8; DB 6; Length 267;
Best Local Similarity 90.0%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 2 gtggatcacaaagattgat 21
    ||||| ||||| |||
Db 180 gtggatgacaaagattgat 199

```

```

RESULT 7
US-09-849-526A-27340/c
: Sequence 27340, Application US/09849526A
: GENERAL INFORMATION:
: APPLICANT: Ruff, Thomas G.
: TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
: FILE REFERENCE: 16517,250138-21(51930)B)
: CURRENT APPLICATION NUMBER: US/09/849,526A
: PRIOR FILING DATE: 2001-05-07
: PRIOR APPLICATION NUMBER: US 60/202,214
: PRIOR FILING DATE: 2000-05-08
: PRIOR APPLICATION NUMBER: US 09/654,617
: PRIOR FILING DATE: 2000-09-05
: PRIOR APPLICATION NUMBER: US 09/684,016
: PRIOR FILING DATE: 2000-10-10
: PRIOR APPLICATION NUMBER: US 09/816,660
: PRIOR FILING DATE: 2001-03-26
: NUMBER OF SEQ ID NOS: 30131
: SEQ ID NO 27340
: LENGTH: 551
: TYPE: DNA
: ORGANISM: Zea mays subsp. mexicana
: FEATURE:
: OTHER INFORMATION: Clone ID: uc-zmflteosinte90d04b1
US-09-849-526A-27340

```

```

Query Match          76.4% Score 16.8; DB 6; Length 551;
Best Local Similarity 90.0%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 3 ttggatcacaaagattgat 22
    ||| ||||| ||||| |||
Db 264 ttggacgacaaagattgat 245

```

```

RESULT 8
US-09-815-264-26466/c
: Sequence 26466, Application US/09815264
: GENERAL INFORMATION:
: APPLICANT: Boukharov, Andrey A.
: APPLICANT: Cao, Yongwei
: APPLICANT: Dotson, Stanton B.
: APPLICANT: Koshi, Jeffrey M.
: APPLICANT: Kovalic, David K.
: APPLICANT: Liu, Jingdong
: APPLICANT: McIninch, James
: APPLICANT: Wu, Wei
: TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof
: FILE REFERENCE: 38-21(51237)G
: CURRENT APPLICATION NUMBER: US/09/815,264
: CURRENT FILING DATE: 2001-03-23
: PRIOR APPLICATION NUMBER: US 09/620,392

```

```

: PRIOR FILING DATE: 2000-07-19
: PRIOR APPLICATION NUMBER: US 09/702,134
: PRIOR FILING DATE: 2000-10-31
: NUMBER OF SEQ ID NOS: 109669
: SEQ ID NO 26466
: LENGTH: 849
: TYPE: DNA
: ORGANISM: Oryza sativa
US-09-815-264-26466

```

```

Query Match          76.4% Score 16.8; DB 6; Length 849;
Best Local Similarity 90.0%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 2 gtggatcacaaagattgat 21
    ||||| ||||| ||||| |||
Db 256 gtggatgacaaagattgat 237

```

```

RESULT 9
US-09-815-264-63319
: Sequence 63319, Application US/09815264
: GENERAL INFORMATION:
: APPLICANT: Boukharov, Andrey A.
: APPLICANT: Cao, Yongwei
: APPLICANT: Dotson, Stanton B.
: APPLICANT: Koshi, Jeffrey M.
: APPLICANT: Kovalic, David K.
: APPLICANT: Liu, Jingdong
: APPLICANT: McIninch, James
: APPLICANT: Wu, Wei
: TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof
: FILE REFERENCE: 38-21(51237)G
: CURRENT APPLICATION NUMBER: US/09/815,264
: PRIOR FILING DATE: 2001-03-23
: PRIOR APPLICATION NUMBER: US 09/620,392
: PRIOR FILING DATE: 2000-07-19
: PRIOR APPLICATION NUMBER: US 09/702,134
: PRIOR FILING DATE: 2000-10-31
: NUMBER OF SEQ ID NOS: 109669
: SEQ ID NO 63319
: LENGTH: 1046
: TYPE: DNA
: ORGANISM: Oryza sativa
US-09-815-264-63319

```

```

Query Match          76.4% Score 16.8; DB 6; Length 1046;
Best Local Similarity 90.0%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 2 gtggatcacaaagattgat 21
    ||||| ||||| ||||| |||
Db 791 gtggatcacaaagattgat 810

```

```

RESULT 10
US-09-815-264-91672
: Sequence 91672, Application US/09815264
: GENERAL INFORMATION:
: APPLICANT: Boukharov, Andrey A.
: APPLICANT: Cao, Yongwei
: APPLICANT: Dotson, Stanton B.
: APPLICANT: Koshi, Jeffrey M.
: APPLICANT: Kovalic, David K.
: APPLICANT: Liu, Jingdong
: APPLICANT: McIninch, James
: APPLICANT: Wu, Wei
: TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof
: FILE REFERENCE: 38-21(51237)G
: CURRENT APPLICATION NUMBER: US/09/815,264
: CURRENT FILING DATE: 2001-03-23

```

```
; PRIOR APPLICATION NUMBER: US 09/620,392
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/702,134
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 109669
; SEQ ID NO 91672
; LENGTH: 1246
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-815-264-91672
```

```
Query Match          76.4%; Score 16.8; DB 6; Length 1246;
Best Local Similarity 90.0%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 2 ttggatcacagaatttgat 21
Db 1013 gtcggactactagattag 1032
```

```
RESULT 11
US-09-815-264-23140
; Sequence 23140, Application US/09815264
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Cao, Yongwei
; APPLICANT: Dotson, Stanton B.
; APPLICANT: Koshl, Jeffrey M.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: McIninch, James
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51237)G
; CURRENT APPLICATION NUMBER: US/09/815,264
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 09/620,392
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/702,134
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 109669
; SEQ ID NO 23140
; LENGTH: 1501
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-815-264-23140
```

```
Query Match          76.4%; Score 16.8; DB 6; Length 1501;
Best Local Similarity 90.0%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 3 ttgatcacagaatttgata 22
Db 674 ttggaccaaagatttgata 693
```

```
RESULT 12
US-09-815-264-46882
; Sequence 46882, Application US/09815264
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Cao, Yongwei
; APPLICANT: Dotson, Stanton B.
; APPLICANT: Koshl, Jeffrey M.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: McIninch, James
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51237)G
; CURRENT APPLICATION NUMBER: US/09/815,264
```

```
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 09/620,392
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/702,134
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 109669
; SEQ ID NO 46882
; LENGTH: 1501
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1501)
; OTHER INFORMATION: unsure at all n locations
US-09-815-264-46882
```

```
Query Match          76.4%; Score 16.8; DB 6; Length 1501;
Best Local Similarity 90.0%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 3 ttgatcacagaatttgata 22
Db 674 ttggaccaaagatttgata 693
```

```
RESULT 13
US-09-815-264-53250
; Sequence 53250, Application US/09815264
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Cao, Yongwei
; APPLICANT: Dotson, Stanton B.
; APPLICANT: Koshl, Jeffrey M.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: McIninch, James
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51237)G
; CURRENT APPLICATION NUMBER: US/09/815,264
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 09/620,392
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/702,134
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 109669
; SEQ ID NO 53250
; LENGTH: 1501
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-815-264-53250
```

```
Query Match          76.4%; Score 16.8; DB 6; Length 1501;
Best Local Similarity 90.0%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 2 gttggatcacagaatttgat 21
Db 88 gtcggactactagattag 107
```

```
RESULT 14
US-09-815-264-76889
; Sequence 76889, Application US/09815264
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Cao, Yongwei
; APPLICANT: Dotson, Stanton B.
; APPLICANT: Koshl, Jeffrey M.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
```

```

; APPLICANT: McIninch, James
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51237)G
; CURRENT APPLICATION NUMBER: US/09/815,264
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 09/620,392
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/702,134
; NUMBER OF SEQ ID NOS: 109669
; SEQ ID NO: 76889
; LENGTH: 6940
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(6940)
; OTHER INFORMATION: unsure at all n locations
US-09-815-264-76889

```

```

Query Match          76.4%; Score 16.8; DB 6; Length 6940;
Best Local Similarity 90.0%; Pred. No. 1.7e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY      3 ttggatcacagaatttgata 22
      ||||| || ||||| ||||| |||||
DB      3165 ttggaccacaaagatttgata 3184

```

```

RESULT 15
US-09-815-264-81045/C
; Sequence 81045, Application US/09815264
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Cao, Yongwei
; APPLICANT: Dotson, Stanton B.
; APPLICANT: Koshl, Jeffrey M.
; APPLICANT: Kovallig, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: McIninch, James
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51237)G
; CURRENT APPLICATION NUMBER: US/09/815,264
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 09/620,392
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/702,134
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 109669
; SEQ ID NO: 81045
; LENGTH: 12081
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-815-264-81045

```

```

Query Match          76.4%; Score 16.8; DB 6; Length 12081;
Best Local Similarity 90.0%; Pred. No. 1.7e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY      2 gttggatcacagaatttgat 21
      ||||| ||||| ||||| |||||
DB      4391 GTTGGATCACTAGATTAGAT 4372

```

Search completed: December 15, 2001, 05:58:26
Job time: 12269 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 03:19:57 ; Search time 2725.73 Seconds
(without alignments)
133.153 Million cell updates/sec

Title: US-09-380-826a-7

Perfect score: 1 tglgtacacagattgata 22

Sequence: IDENTITY_NUC

Scoring table: Gapop 10.0, Gapext 1.0

Searched: 1472140 segs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: gb-da:*
2: gb-hcg:*
3: gb-in:*
4: gb-om:*
5: gb-ov:*
6: gb-pat:*
7: gb-ph:*
8: gb-pl:*
9: gb-pr:*
10: gb-ro:*
11: gb-sts:*
12: gb-sy:*
13: gb-un:*
14: gb-vi:*
15: gb-da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vi:*
30: em_htgo_hum:*
31: em_htgo_inv:*
32: em_htgo_rnd:*
33: em_htg_hum:*
34: em_htg_inv:*
35: em_htg_rnd:*
36: em_htg_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	22	100.0	1481	1	LFU60594	U60594 Leptospira
2	20.4	92.7	354	1	AB007014	AB007014 Spirochae
3	18.8	85.5	283	1	AF003953	AF003953 Leptospir
4	18.8	85.5	288	1	LIU94974	LIU94974 Leptospira
5	18.8	85.5	288	1	LIU94975	LIU94975 Leptospira
6	18.8	85.5	288	1	LIU94976	LIU94976 Leptospira
7	18.8	85.5	288	1	LIU94977	LIU94977 Leptospira
8	18.8	85.5	288	1	LIU94978	LIU94978 Leptospira
9	18.8	85.5	288	1	LIU94979	LIU94979 Leptospira
10	18.8	85.5	353	1	AB007012	AB007012 Spirochae
11	18.8	85.5	1012	8	ATRHNS2X	M98336 Arabidopsis
12	18.8	85.5	9423	8	ATH275979	AT275979 Arabidops
13	18.8	85.5	72590	8	AC003000	AC003000 Arabidops
14	18.8	85.5	121524	2	AF165146	AF165146 Homo sapi
15	18.8	85.5	123090	2	AC083961	AC083961 Homo sapi
16	18.8	85.5	182545	2	AC046176	AC046176 Homo sapi
17	18.8	85.5	200799	8	AF137379	AF137379 Nephrosel
18	18.8	85.5	200799	8	AF137379	AF137379 Nephrosel
19	18.4	83.6	1812	3	AF220067	AF220067 Drosophill
20	18.4	83.6	27365	2	AC015209	AC015209 Drosophill
21	18.4	83.6	173613	3	AC007475	AC007475 Drosophill
22	18.4	83.6	262731	3	AE003823	AE003823 Drosophill
23	17.8	80.9	1874	14	EBBPC4ANCP	EBBPC4ANCP
24	17.8	80.9	31214	8	SPCC830	AL109850 S.pombe c
25	17.8	80.9	85992	2	AB009052	AB009052 Arabidops
26	17.8	80.9	91894	2	AC026327	AC026327 Homo sapi
27	17.8	80.9	172681	9	AC026320	AC026320 Homo sapi
28	17.8	80.9	270889	2	AC055742	AC055742 Homo sapi
29	17.4	79.1	59261	8	T12M4	AC003114 Arabidops
30	17.4	79.1	81662	8	AB008265	AB008265 Arabidops
31	17.4	79.1	109476	9	AL513548	AL513548 Human DNA
32	17.4	79.1	109512	2	AC068144	AC068144 Homo sapi
33	17.4	79.1	148997	9	AC021015	AC021015 Homo sapi
34	17.4	79.1	187847	2	AC023155	AC023155 Homo sapi
35	17.2	78.2	316	1	AF352068	AF352068 Peptostre
36	17.2	78.2	462	1	AF175672	AF175672 Unculture
37	17.2	78.2	835	8	AF271231	AF271231 Albungo ca
38	17.2	78.2	1083	1	AF018567	AF018567 Unidentif
39	17.2	78.2	2213	8	LEP4CCOALB	D49367 Lithospermia
40	17.2	78.2	3608	8	SCYL031C	D73136 S.cerevisia
41	17.2	78.2	5596	3	SPU40832	U40832 Strongyloce
42	17.2	78.2	34496	3	U41016	U41016 Caenorhabdi
43	17.2	78.2	38655	3	U41545	U41545 Caenorhabdi
44	17.2	78.2	49311	2	P26K10	AL049803 Arabidops
45	17.2	78.2	60406	9	AL590550	AL590550 Human DNA

ALIGNMENTS

RESULT 1
LOCUS LFU60594 1481 bp DNA BCT 10-SEP-1998
DEFINITION Leptospira fainei 16S ribosomal RNA gene, partial sequence.
ACCESSION U60594
VERSION U60594.1 GI:1408219
KEYWORDS
SOURCE
ORGANISM Leptospira fainei.
Bacteria; Spirochaetales; Leptospiraceae; Leptospira.
REFERENCE
1 (bases 1 to 1481)
Perolat,P., Chappel,R.J., Adler,B., Baranton,G., Bulach,D.M.,
Billinghurst,M.L., Letocart,M., Merlen,F. and Serrano,M.S.
Leptospira fainei sp. nov., isolated from pigs in Australia
Int. J. Syst. Bacteriol. 48 Pt 3, 851-858 (1998)
TITLE JOURNAL
MEDLINE
REFERENCE
2 (bases 1 to 1481)
Adler,B., Chappel,R.J., Baranton,G., Bulach,D.M.,
Billinghurst,M.L., Letocart,M., Merlen,F., Serrano,M.S. and
Perolat,P.

TITLE Direct Submission
JOURNAL Submitted (12-JUN-1996) Microbiology, Monash University, Wellington
RD., Clayton, VIC 3168, Australia

FEATURES
source
1.1481
/organism="Leptospira fainei"
/strain="Hurstbridge"
/db_xref="taxon:48782"
<1..>1481
/product="16S ribosomal RNA"

BASE COUNT 391 a 335 c 439 g 314 t 2 others

ORIGIN

Query Match 100.0%; Score 22; DB 1; Length 1481;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttgatcacagaattgata 22
|||||
Db 155 TGTGGATCACAGATTGATA 176

RESULT 2
AB007014 354 bp DNA BCT 13-OCT-1997
LOCUS Spirochaeta sp. 16S rRNA gene, partial sequence.
DEFINITION AB007014
AB007014.1 GI:2516255
16S ribosomal RNA.
VERSION Spirochaeta sp. (sub_species:Freshwater obligate oligotroph,
strain:RO-95) DNA.
KEYWORDS Spirochaeta sp.
ORGANISM Spirochaetales; Spirochaetaceae; Spirochaeta.
REFERENCE 1 (bases 1 to 354)
AUTHORS Shin,M.-S.
TITLE Direct Submission
JOURNAL Submitted (05-SEP-1997) to the DDBJ/EMBL/Genbank databases. Mi-Sun
Shin, Laboratory of Marine Molecular Microbiology, Faculty of
Agriculture, Kyoto University, Okwakencho, Kitashirakawa, Sakyo-ku,
Kyoto, Kyoto 606-01, Japan (E-mail:misun@kais.kyoto-u.ac.jp,
Tel:075-753-6224, Fax:075-753-6226)
2 (sites)
REFERENCE Shin,M., Yoshinaga,I., Uchida,A. and Ishida,Y.
AUTHORS Phylogenetic analysis by 16S rRNA gene sequencing of obligate
TITLE oligotrophs isolated from the northern basin of Lake Biwa
(Mesotrophic Lake)
JOURNAL Unpublished (1997)
FEATURES
source
Location/Qualifiers
1..354
/organism="Spirochaeta sp."
/strain="RO-95"
/sub_species="Freshwater obligate oligotroph"
/db_xref="taxon:28185"
<1..>354
rRNA /product="16S rRNA"
BASE COUNT 91 a 79 c 111 g 73 t
ORIGIN

Query Match 92.7%; Score 20.4; DB 1; Length 354;
Best Local Similarity 95.5%; Pred. No. 13;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 tgttgatcacagaattgata 22
|||||
Db 133 TGTGGATCACAGATTGATA 154

RESULT 3
AF003953 283 bp DNA BCT 30-MAY-1998
LOCUS Leptospira interrogans strain 48/95 16S ribosomal RNA gene, partial
DEFINITION

sequence.
ACCESSION AF003953
VERSION AF003953.1 GI:3169306
KEYWORDS
SOURCE
ORGANISM Leptospira interrogans.
REFERENCE 1 (bases 1 to 283)
AUTHORS Woo,T.H.S., Patel,B.K.C., Cinco,M., Smythe,L.D., Symonds,M.,
Norris,M. and Dohnt,M.
TITLE Identification of Leptonema by real-time homogeneous assay of rapid
cycle PCR product
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 283)
Woo,T.H.S., Patel,B.K.C., Cinco,M., Smythe,L.D., Symonds,M.,
Norris,M. and Dohnt,M.
TITLE Direct Submission
JOURNAL Submitted (14-MAY-1997) School of Science, Griffith University,
Brisbane, QLD 4111, Australia
FEATURES
source
Location/Qualifiers
1..283
/organism="Leptospira interrogans"
/strain="48/95"
/db_xref="taxon:173"
<1..>283
rRNA /product="16S rRNA"
BASE COUNT 74 a 60 c 88 g 61 t
ORIGIN

Query Match 85.5%; Score 18.8; DB 1; Length 283;
Best Local Similarity 90.9%; Pred. No. 78;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 tgttgatcacagaattgata 22
|||||
Db 136 TATGGATCACAGATTGATA 157

RESULT 4
L1U94974 288 bp DNA BCT 01-JAN-1998
LOCUS Leptospira inada1 16S ribosomal RNA gene, partial sequence.
DEFINITION U94974
U94974.1 GI:2735446
1 (bases 1 to 288)
REFERENCE Woo,T.H.S., Smythe,L.D., Symonds,M., Norris,M., Dohnt,M.,
AUTHORS Brenner,D.J. and Patel,B.K.C.
TITLE Identification of Leptospira inada1 by continuously monitoring
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 288)
Woo,T.H.S., Smythe,L.D., Symonds,M., Norris,M., Dohnt,M.,
AUTHORS Brenner,D.J. and Patel,B.K.C.
TITLE Direct Submission
JOURNAL Submitted (24-MAR-1997) School of Science, Griffith University,
Brisbane, QLD 4111, Australia
FEATURES
source
Location/Qualifiers
1..288
/organism="Leptospira inada1"
/strain="79/95"
/db_xref="taxon:29506"
<1..>288
rRNA /product="16S ribosomal RNA"
BASE COUNT 77 a 61 c 88 g 62 t
ORIGIN

Query Match 85.5%; Score 18.8; DB 1; Length 288;

Best Local Similarity 90.9%; Pred. No. 78;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 tgttgatcacaaagttagata 22
1 ||||||| |||||||
Db 141 TATTGGATCAGAGATTGTGATA 162

RESULT 5

LIU94975 288 bp DNA BCT 01-JAN-1998
LOCUS Leptospira inadai 16S ribosomal RNA gene, partial sequence.
DEFINITION U94975
ACCESSION U94975
VERSION U94975.1 GI:2735447

KEYWORDS Leptospira inadai.
SOURCE Leptospira inadai.
ORGANISM Bacteria; Spirochaetales; Leptospiraceae; Leptospira.

REFERENCE 1 (bases 1 to 288)
Woo, T.H.S., Smythe, L.D., Symonds, M., Norris, M., Dohnl, M.,

TITLE Brenner, D.J. and Patel, B.K.C.
Identification of Leptospira inadai by continuously monitoring
fluorescence during rapid cycle PCR
Unpublished

JOURNAL 2 (bases 1 to 288)
REFERENCE Woo, T.H.S., Smythe, L.D., Symonds, M., Norris, M., Dohnl, M.,
AUTHORS Brenner, D.J. and Patel, B.K.C.

TITLE Direct Submission
JOURNAL Submitted (24-MAR-1997) School of Science, Griffith University,
Brisbane, QLD 4111, Australia

FEATURES

source 1..288
/organism="Leptospira inadai"
/strain="68/94"
/db_xref="taxon:29506"
rRNA <1..>288
/product="16S ribosomal RNA"
BASE COUNT 77 a 61 c 88 g 62 t
ORIGIN

Query Match 85.5%; Score 18.8; DB 1; Length 288;
Best Local Similarity 90.9%; Pred. No. 78;

Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 tgttgatcacaaagttagata 22
1 ||||||| |||||||
Db 141 TATTGGATCAGAGATTGTGATA 162

RESULT 6

LIU94976 288 bp DNA BCT 01-JAN-1998
LOCUS Leptospira inadai 16S ribosomal RNA gene, partial sequence.
DEFINITION U94976
ACCESSION U94976
VERSION U94976.1 GI:2735448

KEYWORDS Leptospira inadai.
SOURCE Leptospira inadai.
ORGANISM Bacteria; Spirochaetales; Leptospiraceae; Leptospira.

REFERENCE 1 (bases 1 to 288)
Woo, T.H.S., Smythe, L.D., Symonds, M., Norris, M., Dohnl, M.,

TITLE Brenner, D.J. and Patel, B.K.C.
Identification of Leptospira inadai by continuously monitoring
fluorescence during rapid cycle PCR
Unpublished

JOURNAL 2 (bases 1 to 288)
REFERENCE Woo, T.H.S., Smythe, L.D., Symonds, M., Norris, M., Dohnl, M.,
AUTHORS Brenner, D.J. and Patel, B.K.C.

TITLE Direct Submission
JOURNAL Submitted (24-MAR-1997) School of Science, Griffith University,
Brisbane, QLD 4111, Australia

FEATURES

source 1..288
/organism="Leptospira inadai"
/strain="346/95"
/db_xref="taxon:29506"

rRNA <1..>288
/product="16S ribosomal RNA"
BASE COUNT 77 a 61 c 88 g 62 t
ORIGIN

Query Match 85.5%; Score 18.8; DB 1; Length 288;
Best Local Similarity 90.9%; Pred. No. 78;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 tgttgatcacaaagttagata 22
1 ||||||| |||||||
Db 141 TATTGGATCAGAGATTGTGATA 162

RESULT 7

LIU94977 288 bp DNA BCT 01-JAN-1998
LOCUS Leptospira inadai 16S ribosomal RNA gene, partial sequence.
DEFINITION U94977
ACCESSION U94977
VERSION U94977.1 GI:2735449

KEYWORDS Leptospira inadai.
SOURCE Leptospira inadai.

ORGANISM Bacteria; Spirochaetales; Leptospiraceae; Leptospira.
REFERENCE 1 (bases 1 to 288)
Woo, T.H.S., Smythe, L.D., Symonds, M., Norris, M., Dohnl, M.,

TITLE Brenner, D.J. and Patel, B.K.C.
Identification of Leptospira inadai by continuously monitoring
fluorescence during rapid cycle PCR
Unpublished

JOURNAL 2 (bases 1 to 288)
REFERENCE Woo, T.H.S., Smythe, L.D., Symonds, M., Norris, M., Dohnl, M.,
AUTHORS Brenner, D.J. and Patel, B.K.C.

TITLE Direct Submission
JOURNAL Submitted (24-MAR-1997) School of Science, Griffith University,
Brisbane, QLD 4111, Australia

FEATURES

source 1..288
/organism="Leptospira inadai"
/strain="268/95"
/db_xref="taxon:29506"
rRNA <1..>288
/product="16S ribosomal RNA"
BASE COUNT 77 a 61 c 88 g 62 t
ORIGIN

Query Match 85.5%; Score 18.8; DB 1; Length 288;
Best Local Similarity 90.9%; Pred. No. 78;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 tgttgatcacaaagttagata 22
1 ||||||| |||||||
Db 141 TATTGGATCAGAGATTGTGATA 162

RESULT 8

LIU94978 288 bp DNA BCT 01-JAN-1998
LOCUS Leptospira inadai 16S ribosomal RNA gene, partial sequence.
DEFINITION U94978
ACCESSION U94978
VERSION U94978.1 GI:2735450

KEYWORDS Leptospira inadai.
SOURCE Leptospira inadai.
ORGANISM Bacteria; Spirochaetales; Leptospiraceae; Leptospira.

REFERENCE 1 (bases 1 to 288)
Woo, T.H.S., Smythe, L.D., Symonds, M., Norris, M., Dohnl, M.,

TITLE	Brenner,D.J. and Patel,B.K.C.				
JOURNAL	Identification of Leptospira inadai by continuously monitoring fluorescence during rapid cycle PCR				
REFERENCE	Unpublished				
AUTHORS	2 (bases 1 to 288)				
TITLE	Woo,T.H.S., Smythe,L.D., Symonds,M., Norris,M., Dohnt,M., Brenner,D.J. and Patel,B.K.C.				
JOURNAL	Direct Submission				
FEATURES	Submitted (24-MAR-1997) School of Science, Griffith University, Brisbane, QLD 4111, Australia				
SOURCE	Location/Qualifiers				
ORIGIN	1..288				
BASE COUNT	/organism="Leptospira inadai"				
ORIGIN	/strain="218/95"				
BASE COUNT	/db_xref="taxon:29506"				
ORIGIN	<1..>288				
BASE COUNT	/product="16S ribosomal RNA"				
ORIGIN	77 a 61 c 88 g 62 t				
ORIGIN	Query Match				
Best Local Similarity	85.5%; Score 18.8; DB 1; Length 288;				
Matches	20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;				
OY	1 tcttgatcacagaatttgata 22				
Db	141 TATTGGATCACAGATTGGATA 162				
RESULT	9				
LOCUS	L1094979 288 bp DNA BCT 01-JAN-1998				
DEFINITION	Leptospira inadai 16S ribosomal RNA gene, partial sequence.				
ACCESSION	U94979				
VERSION	U94979.1 GI:2735451				
KEYWORDS	.				
SOURCE	Leptospira inadai.				
ORGANISM	Leptospira inadai.				
REFERENCE	Bacteria: Spirochaetales; Leptospiraceae; Leptospira.				
AUTHORS	1 (bases 1 to 288)				
TITLE	Woo,T.H.S., Smythe,L.D., Symonds,M., Norris,M., Dohnt,M., Brenner,D.J. and Patel,B.K.C.				
JOURNAL	Identification of Leptospira inadai by continuously monitoring fluorescence during rapid cycle PCR				
REFERENCE	Unpublished				
AUTHORS	2 (bases 1 to 288)				
TITLE	Woo,T.H.S., Smythe,L.D., Symonds,M., Norris,M., Dohnt,M., Brenner,D.J. and Patel,B.K.C.				
JOURNAL	Direct Submission				
FEATURES	Submitted (24-MAR-1997) School of Science, Griffith University, Brisbane, QLD 4111, Australia				
SOURCE	Location/Qualifiers				
ORIGIN	1..288				
BASE COUNT	/organism="Leptospira inadai"				
ORIGIN	/strain="1078 VRI"				
BASE COUNT	/db_xref="taxon:29506"				
ORIGIN	<1..>288				
BASE COUNT	/product="16S ribosomal RNA"				
ORIGIN	77 a 61 c 88 g 62 t				
ORIGIN	Query Match				
Best Local Similarity	85.5%; Score 18.8; DB 1; Length 288;				
Matches	20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;				
OY	1 tcttgatcacagaatttgata 22				
Db	141 TATTGGATCACAGATTGGATA 162				
RESULT	10				

LOCUS	AB007012	353 bp	DNA	BCF	13-OCT-1997
DEFINITION	Spirochaeta sp. 16S rRNA gene, partial sequence.				
ACCESSION	AB007012				
VERSION	AB007012.1	GI:2516253			
KEYWORDS	16S ribosomal RNA.				
SOURCE	Spirochaeta sp. (sub_species:Freshwater obligate oligotroph, strain:SO-104) DNA.				
ORGANISM	Spirochaeta sp.				
REFERENCE	Bacteria: Spirochaetales; Spirochaetaceae; Spirochaeta.				
AUTHORS	1 (bases 1 to 353)				
TITLE	Shin,M., S.				
REFERENCE	Direct Submission				
AUTHORS	Submitted (05-SEP-1997) to the DDBJ/EMBL/GenBank databases. Mi-Sun Shin, Laboratory of Marine Molecular Microbiology, Faculty of Agriculture, Kyoto University; Oiwakecho, Kitashirakawa, Sakyo-ku, Kyoto, Kyoto 306-01, Japan (E-mail:mtsun@kais.kyoto-u.ac.jp, Tel:075-753-6224, Fax:075-753-6226)				
TITLE	2 (sites)				
REFERENCE	Shin,M., Yoshinaga,I., Uchida,A. and Ishida,Y.				
AUTHORS	Phylogenetic analysis by 16S rRNA gene sequencing of Obligate				
TITLE	Oligotrophic Isolated from the northern basin of Lake Biwa				
REFERENCE	(Mesotrophic Lake)				
FEATURES	Unpublished (1997)				
SOURCE	Location/Qualifiers				
	1..353				
	/organism="Spirochaeta sp."				
	/strain="SO-104"				
	/sub_species="Freshwater obligate oligotroph"				
	/db_xref="taxon:28185"				
	<1..>353				
	/product="16S rRNA"				
BASE COUNT	90 a	80 c	108 g	75 t	
ORIGIN					
Query Match	85.5%;	Score 18.8;	DB 1;	Length 353;	
Best Local Similarity	90.9%;	Pred. NO. 77;			
Matches	20; Conservative	0; Mismatches	2; Indels	0; Gaps	0.
OY	1	tggtgattcacagaatttgata	22		
Db	132	TGTTGGATCACCAGATCTGATA	153		
RESULT 11					
ATHRNS2X	1012 bp	mRNA	PLN	30-OCT-1994	
LOCUS	Arabidopsis thaliana ribonuclease (RNS2) mRNA, complete cds.				
DEFINITION	M98336				
ACCESSION	M98336.1				
VERSION	GI:289209				
KEYWORDS	ribonuclease.				
SOURCE	Arabidopsis thaliana cDNA to mRNA.				
ORGANISM	Arabidopsis thaliana				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;				
	Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.				
	1 (bases 1 to 1012)				
REFERENCE	Taylor,C.B., Bartola,P.A., delCardayre,S.B., Raines,R.T. and				
AUTHORS	Green,P.J.				
TITLE	RNS2: a senescence-associated RNase of Arabidopsis that diverged				
JOURNAL	from the S-RNases before speciation				
MEDLINE	Proc. Natl. Acad. Sci. U.S.A. 90 (11), 5118-5122 (1993)				
FEATURES	93281708				
Source	Location/Qualifiers				
	1..1012				
	/organism="Arabidopsis thaliana"				
	/db_xref="taxon:3702"				
	16..795				
gene	/gene="RNS2"				
	16..795				
CDS	/gene="RNS2"				
	/codon_start=1				

Genes were identified by a combination of three methods: Gene prediction programs including GRAIL (<http://arthur.epm.ornl.gov/pub/xtgrail>), GeneFinder (Phil Green, University of Washington), Genscan (Chris Burge, <http://genome.stanford.edu/GENSCAN.html>), and NetPlantGene (<http://www.cds.dtu.dk/services/NetGene2/>), searches of the complete sequence against a peptide database and plant EST databases at TIGR, and manual curation based on those analyses. Annotated genes are named to indicate the level of evidence for their annotation. Genes with similarity to other proteins are named after the database hits. Genes without significant peptide similarity but with EST similarity are named as 'unknown' proteins. Genes without protein or EST similarity, that are predicted by two or more gene prediction programs over most of their length are annotated as 'hypothetical' proteins. Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, <http://genome.wustl.edu/eddy/tRNAscan-SE/>). Simple repeats were identified by RepeatMasker (Arian Smit, <http://ftp.genome.washington.edu/RM/RepeatMasker.html>). Genes are numbered from the top to bottom of the chromosome.

We thank the CSHL/MashU/ABI consortium for sequencing BAC clones FEP23, F5J6, T17A5, and T13J16, the ESSA group for sequencing clone F13J4, and Scott Jackson, Jiming Jiang, Klaus Meyer, Eric Richards and Satoshi Tabata for helpful assistance. In addition, we would like to thank the TIGR Bioinformatics Department, especially Lixin Zhou, Hanli Khalak, Michael E. Heaney, Lily Fu, Feng Liang, Jeremy Peterson, Michael Holmes, and Delwood Richardson for software and database support.

This work was supported by the National Science Foundation, Department of Energy and the US Department of Agriculture.

Address all correspondence to: ateligr.org.

FEATURES		FEATURES	
source	Location/Qualifiers	source	Location/Qualifiers
misc_feature	1..72590 /organism="Arabidopsis thaliana" /cultivar="Columbia" /db_xref="taxon:3702" /chromosome="11" 1..>72590 /note="Sequence from clone T517"	gene	complement(<10039..11304) /gene="At2g39720" /note="T517.2" complement(10039..11244) /gene="At2g39720" /note="unknown protein" /codon_start=1 /protein_id="AAB87121.1" /db_xref="GI:2642154"
repeat_region	/translating="MDSGSSSSLSKFNLSIVLLIFLPTPKSTSTNOLLSLK TOLKPOSSDKLSFRHNVTLTLVTLAVGPPONIEMVLDTGSELSMLCKKSPNLSVY NPVSSSTYSPVSGSPICRTTRDLPIASCDPTLCHVAISYADATISIGNLAHET FVTSVTPRPTLRCGMSGCSNSEEDAKSGILMGNGSLSFVNOIGFESKESYCSIG SDSSGFLLLGDASTSWLGPITYPLVLOSITPLPFDRAATVQLEGIRVSKILSLK SVFVDDHTGAGTQVWDSGTQFTLMGPVYTLKNEFTOTKSVLRLVDDPVCFOGT DLCKVSGTTPNFSGILPMSILMRGAMSVSGKLLYRVNGASGEDEEYVCTFGN SDLGIEAFVIGHHQQNVMMFEDLAKSRVGFAGNVRCDLASORLGLND"	CDS	complement(11453..11512) /rpt_family="POLY_A" complement(join(13323..13599,13724..14005,14087..14559, 14651..14736,14836..14941,15033..15349,15834..16055)) /gene="At2g39730"
repeat_region	/rpt_family="POLY_A" complement(11266..1291) /rpt_family="(CAAA)n"	mRNA	complement(join(13323..13599,13713..14005,14087..14559, 14651..14736,14836..14941,15033..15349,15834..16055)) /gene="At2g39730"
repeat_region	/rpt_family="POLY_A" complement(2759..2862) /rpt_family="POLY_A" complement(<4584..>5909) /gene="At2g39710"	gene	complement(13323..16055) /gene="At2g39730"
mRNA	complement(<4584..>5909) /gene="At2g39710"	repeat_region	complement(13364..13461) /rpt_family="POLY_A" complement(join(13484..13599,13724..14005,14087..14559, 14651..14736,14836..14941,15033..15349,15834..15878)) /note="alternative splice form 1 GP1166834IM86720"
gene	complement(<4584..>5909) /gene="At2g39710"	CDS	/codon_start=1 /product="Rubisco activase" /protein_id="AAB87122.1" /db_xref="GI:2642155"
CDS	/note="T517.1: similar to GP122450121gn1 PID e3270001 297341" complement(4584..5909) /gene="At2g39710"	repeat_region	/translating="MAAASVTGAINRAPLSLNGSGGSAVASAPSTFLCKKYYVSFR AOSNKSNGSFKYLAVEDKOTGDNRRLGATYPTSDQOQITRGKMGVSVQAPGCT GTHAAVLSYSEYVSGLRQYNLDNMMDGFTIAPFMKLVVHTTKNFTLPNTKVP LGIMGSGGKGSFQCELVMAKGINPTMNSAGLESGNAGEPAKLIROYREAAADLIK KGWCKCLFINDLDAGAGRMGCTQYTNVNMVATLMN1ADNPTNQLPGMYKKEENA RVPTICTGNDFSTLYAPLIRDRGMEKEFYAAPREDRITGVCKGIFRTDKIKDEDIVTLV DPPGOSIDFFGALRARVYVDDEVKRPESIGVEKIGRLNSREGPVDFQPMPTYEK LMEYGNLWMEQENKRVQALAEYTLSSAALGDANADAIGKGTGYKTEKEPEPK"
mRNA	/note="unknown protein" /codon_start=1 /protein_id="AAB87120.1" /db_xref="GI:2642153"	gene	complement(11761..17639) /rpt_family="(TA)n" join(<18783..18921,19188..19352,19496..19674,19766..19861, 19965..20034,20131..20261,20593..>21237) /gene="At2g39740"
mRNA	/note="unknown protein" /codon_start=1 /protein_id="AAB87120.1" /db_xref="GI:2642153"	CDS	<18783..>21237 /gene="At2g39740" /note="T517.4: predicted by genscan" join(18783..18921,19188..19352,19496..19674,19766..19861, 19965..20034,20131..20261,20593..21237)

```

/gene="At2g39740"
/notes="hypothetical protein"
/codon_start=1
/protein_id="AAB67123.1"
/db_xref="gi:2642156"
/translation="MSRNPFLDPTLQELIIVIKPPARDRDTITVIDLRVLSVSEC
LRGATVDFGFSVNLFTRMGDLISVDFGSSSLTFGKKOKOTLGLRLARASG
LWKLOFVTHARVILKAVSGHORISCDISDNLGLSKSELMFISIDGRFRLVLT
LVKEMAKAHNINDSKTGFENYSLSLIVHPFCVAPLLEPLVIYIKRSVADLTG
KRTAEESIAQVTTANINIAFKSEKRSKSVNRSSLVSEFAKVEDPEQPVNAARSY
RRLNDRIAQVFIITSRILVSECNRSNLSIGLITGQHIQESLIRTLSPQIHANMHN
RNLGQARPOQMOQOMNSOSYNTPNPMPPLTOSRPOQMTONTNPNRNLGQPPVQ
QTPVITOTQOQSPYKSGNRPLKNTSAGSSQNGHIGKPSGHMNGVNSARPAYTG
VNSARPSKIPISOGQIWRPRHEO"
complement(join(21380..21726,21839..21916,22124..22246,
22328..22585,22834..22993,23212..23322,23481..23751,
23847..23996,24089..>24814))
/gene="At2g39750"
complement(21380..>24814)

```

```

Query Match      85.5%; Score 18.8; DB 8; Length 72590;
Best Local Similarity 90.9%; Pred. No. 72;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Oy 1 tcttgatcacagaattgata 22
||||| ||||||| |||||
Db 35807 TGTGTGTCACAGATTGTCACA 35828

```

```

RESULT 14
AF165146 121524 bp DNA HTG 07-JUN-2001
LOCUS Homo sapiens chromosome 8 clone CTA-397H3 map 8q12-8q13, ***
DEFINITION SEQUENCING IN PROGRESS ***, 7 unordered pieces.
ACCESSION AF165146
VERSION AF165146.3 GI:14327840
KEYWORDS HTG; HTGS_PHASE1; HTGS_FULLTOP.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

```

REFERENCE 1 (bases 1 to 121524)
AUTHORS Schillabel,M.B., Baumgart,C., Blechschmidt,K., Dette,M., Jahn,N.,
Lehmann,R., Menzel,U., Polley,A., Reichwald,K., Schudy,A.,
Siddiqui,R., Taudien,S., Wen,G., Schlegelberger,B., Siebert,R.,
Rosenthal,A. and Platzer,M.
Chromosome 8 genomic sequence
Unpublished
2 (bases 1 to 121524)
Schudy,A., Blechschmidt,K., Schillabel,M., Baumgart,C., Menzel,U.,
Weber,J., Schattevoy,R. and Rosenthal,A.
Direct Submission
Submitted (06-JUL-1999) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
On Jun 7, 2001 this sequence version replaced gi:8151945.
----- Genome Center
Center: Institute of Molecular Biotechnology
Center code: IMB
Web site: http://genome.imb-jena.de/
Contact: gscj-submit@genome.imb-jena.de
----- Project Information
Center project name: H171
Center clone name: CTA-397H3
----- Summary Statistics
Sequencing Vector: M13; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 115981 bases at least Q40
Consensus quality: 117981 bases at least Q30
Consensus quality: 119332 bases at least Q20
Quality coverage: 9.57 in Q20 bases; sum-of-contigs
-----

```

* NOTE: This is a 'working draft' sequence. It currently

```

* consists of 7 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 1187: contig of 1187 bp in length
1188 1287: gap of unknown length
1288 3530: contig of 2243 bp in length
1288 3531: gap of unknown length
3630: gap of unknown length
21624: contig of 17994 bp in length
21625 21724: gap of unknown length
21725 43656: contig of 21932 bp in length
43657 43757: gap of unknown length
43757 61383: contig of 17627 bp in length
61384 61483: gap of unknown length
61484 85318: contig of 23835 bp in length
85319 85418: gap of unknown length
85419 121524: contig of 36106 bp in length.
Location/Qualifiers
1..121524
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/map="8q12-8q13"
/clone="CTA-397H3"

```

```

BASE COUNT 34959 a 25555 c 25579 g 34831 t 600 others
ORIGIN

```

```

Query Match      85.5%; Score 18.8; DB 2; Length 121524;
Best Local Similarity 90.9%; Pred. No. 71;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Oy 1 tcttgatcacagaattgata 22
||||| ||||||| |||||
Db 7103 TTTTGATCACAAGATTGCTGA 7124

```

```

RESULT 15
AC083961 129090 bp DNA HTG 15-FEB-2001
LOCUS Homo sapiens chromosome 8 clone RP11-22E14 map 8, *** SEQUENCING IN
DEFINITION PROGRESS ***, 67 unordered pieces.
ACCESSION AC083961
VERSION AC083961.2 GI:12831382
KEYWORDS HTG; HTGS_PHASE1.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 129090)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Barra,N., Bastien,V., Beda,F., Boguslavsky,L.,
Bouckgalter,B., Brown,A., Burkett,G., Campoliano,A., Castle,A.,
Chapel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
DeBellando,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,
Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M.,
Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L.,
Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., LaRoque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,C., Liu,G.,
Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McKernan,K.,
McPheeters,R., Meldrum,J., Meneus,L., Mihova,T., Mlenga,V.,
Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T.,
O'Donnell,P., O'Neill,D., Olivarez,T.M., Oliver,J., Peterson,K.,
Pierre,N., Pisanl,C., Pollara,V., Raymond,C., Rieback,M., Riley,R.,
Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P.,
Sougnuez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,

```

TITLE
JOURNAL
COMMENT

Strauss, N., Subramanian, A., Talamas, J., Testfaye, S., Theodore, J.,
Tirrell, A., Travers, M., Trigglio, J., Vassiliev, H., Viel, R., Vo, A.,
Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zalnoun, J.,
Zimmer, A., and Zody, M.

Direct Submission

Submitted (08-Oct-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Feb 15, 2001 this sequence version replaced g1:10717228.
All repeats were identified using RepeatMasker:
Smil, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence-submissions@genome.wi.mit.edu

Project Information

Center project name: L11257

Center clone name: 22_E_14

NOTE: This is a 'working draft' sequence. It currently
consists of 67 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

1 1369: contig of 1369 bp in length
* 1370 1469: gap of 100 bp
* 1470 2279: contig of 810 bp in length
* 2280 2379: gap of 100 bp
* 2380 3192: contig of 813 bp in length
* 3193 3292: gap of 100 bp
* 3293 3539: contig of 247 bp in length
* 3540 3639: gap of 100 bp
* 3640 4530: contig of 891 bp in length
* 4531 4630: gap of 100 bp
* 4631 5236: contig of 606 bp in length
* 5237 5336: gap of 100 bp
* 5337 6055: contig of 719 bp in length
* 6056 6155: gap of 100 bp
* 6156 6653: contig of 498 bp in length
* 6654 6753: gap of 100 bp
* 6754 7388: contig of 635 bp in length
* 7389 7488: gap of 100 bp
* 7489 8145: contig of 657 bp in length
* 8146 8245: gap of 100 bp
* 8246 9151: contig of 906 bp in length
* 9152 9251: gap of 100 bp
* 9252 9962: contig of 711 bp in length
* 9963 10062: gap of 100 bp
* 10063 10776: contig of 714 bp in length
* 10777 10876: gap of 100 bp
* 10877 11474: contig of 598 bp in length
* 11475 11574: gap of 100 bp
* 11575 12639: contig of 1065 bp in length
* 12640 12739: gap of 100 bp
* 12740 13952: contig of 1213 bp in length
* 13953 14052: gap of 100 bp
* 14053 14838: contig of 786 bp in length
* 14839 14938: gap of 100 bp
* 14939 15849: contig of 911 bp in length
* 15850 15949: gap of 100 bp
* 15950 18036: contig of 2087 bp in length
* 18037 18135: gap of 100 bp
* 18137 19282: contig of 1146 bp in length
* 19283 19382: gap of 100 bp
* 19383 20437: contig of 1055 bp in length
* 20438 20537: gap of 100 bp
* 20538 21159: contig of 622 bp in length
* 21160 21259: gap of 100 bp
* 21260 22494: contig of 1235 bp in length

22495 22594: gap of 100 bp
* 22595 23478: contig of 884 bp in length
* 23479 23578: gap of 100 bp
* 23579 24600: contig of 1022 bp in length
* 24601 24700: gap of 100 bp
* 24701 24891: contig of 191 bp in length
* 24892 24991: gap of 100 bp
* 24992 26183: contig of 1192 bp in length
* 26184 26283: gap of 100 bp
* 26284 27862: contig of 1579 bp in length
* 27863 27962: gap of 100 bp
* 27963 28581: contig of 619 bp in length
* 28582 28681: gap of 100 bp
* 28682 29511: contig of 830 bp in length
* 29512 29611: gap of 100 bp
* 29612 31038: contig of 1427 bp in length
* 31039 31138: gap of 100 bp
* 31139 32582: contig of 1444 bp in length
* 32583 32682: gap of 100 bp
* 32683 33540: contig of 858 bp in length
* 33541 33640: gap of 100 bp
* 33641 35263: contig of 1623 bp in length
* 35264 35363: gap of 100 bp
* 35364 37036: contig of 1673 bp in length
* 37037 37136: gap of 100 bp
* 37137 38963: contig of 1827 bp in length
* 38964 39063: gap of 100 bp
* 39064 40807: contig of 1744 bp in length
* 40808 40907: gap of 100 bp
* 40908 42896: contig of 1989 bp in length
* 42897 42996: gap of 100 bp
* 42997 44598: contig of 1602 bp in length
* 44599 44698: gap of 100 bp
* 44699 46343: contig of 1645 bp in length
* 46344 46443: gap of 100 bp
* 46444 47574: contig of 1131 bp in length
* 47575 47674: gap of 100 bp
* 47675 50217: contig of 2543 bp in length
* 50218 50317: gap of 100 bp
* 50318 51338: contig of 1021 bp in length
* 51339 51438: gap of 100 bp
* 51439 53743: contig of 2305 bp in length
* 53744 53843: gap of 100 bp
* 53844 55260: contig of 1417 bp in length
* 55261 55360: gap of 100 bp
* 55361 57843: contig of 2483 bp in length
* 57844 57943: gap of 100 bp
* 57944 59865: contig of 1922 bp in length
* 59866 59965: gap of 100 bp
* 59966 62260: contig of 2295 bp in length
* 62261 62360: gap of 100 bp
* 62361 64120: contig of 1760 bp in length
* 64121 64220: gap of 100 bp
* 64221 65979: contig of 1759 bp in length
* 65980 66079: gap of 100 bp
* 66080 68712: contig of 2633 bp in length
* 68713 68812: gap of 100 bp
* 68813 71145: contig of 2333 bp in length
* 71146 71245: gap of 100 bp
* 71246 74155: contig of 2910 bp in length
* 74156 74255: gap of 100 bp
* 74256 77041: contig of 2786 bp in length
* 77042 77141: gap of 100 bp
* 77142 80685: contig of 3544 bp in length
* 80686 80785: gap of 100 bp
* 80786 83825: contig of 3040 bp in length
* 83826 83925: gap of 100 bp
* 83926 87721: contig of 3796 bp in length
* 87722 87821: gap of 100 bp
* 87822 92728: contig of 4907 bp in length
* 92729 92828: gap of 100 bp
* 92829 96652: contig of 3824 bp in length
* 96653 96752: gap of 100 bp

```

* 96753 99906: contig of 3154 bp in length
* 99907 100006: gap of 100 bp
* 100007 102910: contig of 2904 bp in length
* 102911 103010: gap of 100 bp
* 103011 107188: contig of 4178 bp in length
* 107189 107288: gap of 100 bp
* 107289 111047: contig of 3759 bp in length
* 111048 111147: gap of 100 bp
* 111148 116741: contig of 5594 bp in length
* 116742 116841: gap of 100 bp
* 116842 121990: contig of 5149 bp in length
* 121991 122090: gap of 100 bp
* 122091 126622: contig of 4532 bp in length
* 126623 126722: gap of 100 bp
* 126723 129090: contig of 2368 bp in length.

```

FEATURES

```

Source
1. 129090
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/map="8"
/clone="RP11-22E14"
/clone_lib="RPCT-11 Human Male BAC"

```

```

Query Match      85.5%; Score 18.8; DB 2; Length 129090;
Best Local Similarity 90.9%; Pred. No. 71;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 tggtagatcaagaattgata 22
1 | | | | | | | | | | | | | | | | | | | |
Db 6294 TTTTGATCACAAGATTGTA 6315

```

Search completed: December 15, 2001, 03:20:02
Job time: 7692 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 03:29:46 ; Search time 401.91 seconds
(without alignments)
46.929 Million cell updates/sec

Title: US-09-380-826a-7

Perfect score: 22

Sequence: 1 tgttgatcacacagattgata 22

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: N.Geneseq_1101.*
2: /SIDS2/gcgdata/geneseq/geneseq/NA1980.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq/NA1981.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq/NA1982.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq/NA1983.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq/NA1984.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq/NA1985.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq/NA1986.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq/NA1987.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq/NA1988.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq/NA1989.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq/NA1990.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq/NA1991.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq/NA1992.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq/NA1993.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq/NA1996.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	100.0	22	19	AAV58901
2	22	100.0	22	19	AAV58897
3	22	100.0	1477	19	AAV58896
4	18.8	85.5	1012	21	AAV59365
5	18.8	85.5	1069	21	AAV59365
6	17.8	80.9	2084	21	AAV79696
7	17.2	78.2	447	21	AAV8114
8	17.2	78.2	1290	21	AAV42365
9	17.2	78.2	1413	21	AAV46258
10	17.2	78.2	2038	22	AAV22799
11	16.8	76.4	1764	22	AAH15352

C	12	16.8	76.4	2582	18	AAT90543
C	13	16.4	74.5	2444	21	AAC56059
C	14	16.2	73.6	335	22	AAH33702
C	15	16.2	73.6	951	20	AAV84598
C	16	16.2	73.6	1250	21	AAC59054
C	17	16.2	73.6	1291	19	AAV59636
C	18	16.2	73.6	1376	22	AAH33007
C	19	16.2	73.6	1450	21	AAC49931
C	20	16.2	73.6	1452	21	AAC39657
C	21	16.2	73.6	1473	20	AAH61757
C	22	16.2	73.6	2000	17	AAT71111
C	23	16.2	73.6	2001	17	AAT39904
C	24	16.2	73.6	2001	21	AAZ49816
C	25	16.2	73.6	3822	18	AAT64683
C	26	16.2	73.6	4346	14	AAO42933
C	27	16.2	73.6	5400	21	AAH38732
C	28	16.2	73.6	7470	17	AAT33905
C	29	16.2	73.6	7470	19	AAV29317
C	30	16.2	73.6	7470	21	AAZ49820
C	31	16.2	73.6	10811	19	AAV59091
C	32	15.8	71.8	467	22	AAH32060
C	33	15.8	71.8	670	21	AAC09187
C	34	15.8	71.8	736	21	AAC10808
C	35	15.8	71.8	740	21	AAC51908
C	36	15.8	71.8	843	20	AAZ15467
C	37	15.8	71.8	961	19	AAH14095
C	38	15.8	71.8	4146	21	AAZ51556
C	39	15.6	70.9	297	20	AAV88753
C	40	15.6	70.9	342	22	AAH32175
C	41	15.6	70.9	452	22	AAH31651
C	42	15.6	70.9	635	22	AAH31988
C	43	15.6	70.9	795	20	AAV72025
C	44	15.6	70.9	834	20	AAV72026
C	45	15.6	70.9	1075	18	AAT72715

ALIGNMENTS

RESULT 1	
AAV58901	standard; DNA; 22 BP.
ID	AAV58901
AC	AAV58901;
XX	
DT	20-JAN-1999 (first entry)
XX	
DE	Leptospira RNA gene nucleotide sequence.
XX	
KW	Infection; pathogenic Leptospira; protective immunity; therapy;
KW	diagnosis; ss.
XX	
OS	Leptospira sp.
XX	
PN	W09840099-A1.
XX	
PD	17-SEP-1998.
XX	
PF	06-MAR-1998; 98WO-AU00145.
XX	
PR	07-MAR-1997; 97AU-0005494.
XX	
PA	(AGRI-) AGRIC VICTORIA SERVICES PTY LTD.
PA	(PIGR-) PIG RES & DEV CORP.
XX	
PI	Chappel RJ;
XX	
DR	WPI; 1998-520791/44.
XX	
PT	New isolated pathogenic Leptospira bacterium - useful for, e.g
PT	developing products for conferring protective immunity, and for
PT	prophylactic or therapeutic treatment
XX	

Maize male fertill
Eucalyptus grandis
Human colon cancer
Human secreted pro
Human secreted pro
Human secreted pro
Human colon cancer
Arabidopsis thalia
Arabidopsis thalia
B. burgdorferi ant
Acetyl-CoA-carboxy
Maize acetyl CoA c
ECORI fragment of
M. leprae gyrA pre
A3 maize ACCase cd
Potato subclone pg
Maize acetyl CoA c
Maize ACCase enzym
Maize acetyl CoA c
Potato pollen cell
Human olfactory re
Human secreted pro
Human secreted pro
Arabidopsis thalia
Human gene express
H. pylori GHP0 131
Human hypoxia resp
EST clone HK650.
Human olfactory re
S. epidermidis ope
Human olfactory re
Adenovirus PACTSG2
Adenovirus SCAR-RG
C. elegans inhibit

```
PS Claim 15; Page 72; 94pp; English.
XX
CC This sequence represents a Leptospira DNA sequence isolated from the
CC pathogenic Leptospira (LS) bacterium of the invention. The bacterium
CC belongs to serogroup Hurstbridge or serovar hurstbridge or the species
CC L. fainel. The LS bacteria can be used for conferring protective
CC immunity against pathogenic LS bacteria in humans or animals. The
CC bacteria can also be used for prophylactic or therapeutic treatment of LS
CC infections. The DNAs and antibodies may also be used for detection and
CC diagnosis of past or present LS infection.
XX
SQ Sequence 22 BP; 7 A; 2 C; 5 G; 8 T; 0 other;

Query Match      100.0%; Score 22; DB 19; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tgttgatcacagaattgata 22
   |||||||||||||||||||
Db 1 tgttgatcacagaattgata 22

RESULT 2
AAV58897
ID AAV58897 standard; DNA; 22 BP.
XX
AC AAV58897;
XX
DT 20-JAN-1999 (first entry)
XX
DE L. fainel nucleotide sequence.
XX
KW Infection; pathogenic Leptospira; protective immunity; therapy;
XX diagnosis; ss.
XX
KM Leptospira fainel.
XX
OS
XX
PN WO9840099-A1.
XX
PD 17-SEP-1998.
XX
PF 06-MAR-1998; 98WO-AU00145.
XX
PR 07-MAR-1997; 97AU-0005494.
XX
PA (AGRI-) AGRIC VICTORIA SERVICES PTY LTD.
XX (PIGR-) PIG RES & DEV CORP.
XX
PI Chappel RJ;
XX
DR WPI; 1998-520791/44.
XX
PT New isolated pathogenic Leptospira bacterium - useful for, e.g
PT developing products for conferring protective immunity, and for
PT prophylactic or therapeutic treatment
XX
PS Claim 15; Page 70; 94pp; English.
XX
CC This sequence represents a Leptospira DNA sequence isolated from the
CC pathogenic Leptospira (LS) bacterium of the invention. The bacterium
CC belongs to serogroup Hurstbridge or serovar hurstbridge or the species
CC L. fainel. The LS bacteria can be used for conferring protective
CC immunity against pathogenic LS bacteria in humans or animals. The
CC bacteria can also be used for prophylactic or therapeutic treatment of LS
CC infections. The DNAs and antibodies may also be used for detection and
CC diagnosis of past or present LS infection.
XX
SQ Sequence 22 BP; 7 A; 2 C; 5 G; 8 T; 0 other;

Query Match      100.0%; Score 22; DB 19; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tgttgatcacagaattgata 22
   |||||||||||||||||||
Db 1 tgttgatcacagaattgata 22

RESULT 3
AAV58896
ID AAV58896 standard; DNA; 1477 BP.
XX
AC AAV58896;
XX
DT 20-JAN-1999 (first entry)
XX
DE L. fainel nucleotide sequence.
XX
KW Infection; pathogenic Leptospira; protective immunity; therapy;
XX diagnosis; ss.
XX
KM Leptospira fainel.
XX
OS
XX
PN WO9840099-A1.
XX
PD 17-SEP-1998.
XX
PF 06-MAR-1998; 98WO-AU00145.
XX
PR 07-MAR-1997; 97AU-0005494.
XX
PA (AGRI-) AGRIC VICTORIA SERVICES PTY LTD.
XX (PIGR-) PIG RES & DEV CORP.
XX
PI Chappel RJ;
XX
DR WPI; 1998-520791/44.
XX
PT New isolated pathogenic Leptospira bacterium - useful for, e.g
PT developing products for conferring protective immunity, and for
PT prophylactic or therapeutic treatment
XX
PS Claim 15; Page 69-70; 94pp; English.
XX
CC This sequence represents a Leptospira DNA sequence isolated from the
CC pathogenic Leptospira (LS) bacterium of the invention. The bacterium
CC belongs to serogroup Hurstbridge or serovar hurstbridge or the species
CC L. fainel. The LS bacteria can be used for conferring protective
CC immunity against pathogenic LS bacteria in humans or animals. The
CC bacteria can also be used for prophylactic or therapeutic treatment of LS
CC infections. The DNAs and antibodies may also be used for detection and
CC diagnosis of past or present LS infection.
XX
SQ Sequence 1477 BP; 390 A; 334 C; 439 G; 314 T; 0 other;

Query Match      100.0%; Score 22; DB 19; Length 1477;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tgttgatcacagaattgata 22
   |||||||||||||||||||
Db 154 tgttgatcacagaattgata 175

RESULT 4
AAA99365
ID AAA99365 standard; DNA; 1012 BP.
XX
AC AAA99365;
XX
DT 22-JAN-2001 (first entry)
XX
DE Plant PrG1 promoter related gene sequence.
```

XX Plant promoter; PRAG1; reproductive tissue; transgenic plant; cereal; ds.
KW Arabidopsis thaliana.
OS Arabidopsis thaliana.
XX WO20005172-A1.
PD 21-SEP-2000.
XX 17-MAR-2000; 2000WO-NZ00031.
XX 17-MAR-1999; 99NZ-0334715.
XX (CART-) CARTER HOLT HARVEY LTD.
PA (TASMAN BIOTECHNOLOGY LTD.
PA (UNMKT) UNIV MICHIGAN TECHNOLOGICAL.
XX Podilla GK, Liu J, Karnosky DF;
PI WPI: 2000-594442/56.
DR P-PSDB; AAB26796.
XX Novel plant reproductive tissue promoter, useful to produce plants
PT which have a diminished reproductive capacity or which are sterile
XX Claim 15; Page 42-43; 51pp; English.
XX This invention relates to a novel plant promoter gene. The promoter is
CC located in plant reproductive tissue, and the invention includes
CC transgenic plants containing the promoter. The promoter can be used to
CC produce plants which have a diminished reproductive capacity or which are
CC sterile. The constructs can also be used to transform agronomically
CC important plants in which modulation of reproductive capacity
CC (particularly the timing and abundance of flowering) is desirable,
CC e.g. cereals, rice, maize, wheat, barley, oats, rye, soybean and canola.
CC The present sequence represents DNA encoding the plant reproductive
CC promoter (PRAG1) of the invention.
XX Sequence 1012 BP; 287 A; 197 C; 217 G; 311 T; 0 other;
SQ

Query Match 85.5%; Score 18.8; DB 21; Length 1012;
Best Local Similarity 90.9%; Pred. No. 5.8;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Oy 1 ttttgatcacaaagttagata 22
||||| ||||||||| 4
Db 672 tgttgatcacaaagttagata 693

RESULT 5
AAC37325
ID AAC37325 standard; DNA: 1069 BP.
XX AAC37325;
AC 17-OCT-2000 (first entry)
DT 17-OCT-2000 (first entry)
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 16971.
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 16971.
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX Arabidopsis thaliana.
OS Arabidopsis thaliana.
XX EP1033405-A2.
PN EP1033405-A2.
PD 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
PF 25-FEB-1999; 99US-0121825.
XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145213.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156506.
PR 04-OCT-1999; 99US-0157117.

PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 85.5%; Score 18.8; DB 21; Length 1069;
Best Local Similarity 90.3%; Pred. No. 5.8;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 tgttgatcaccaagatttgata 22
||||| |||||||
Db 690 tgttgatcaccaagatttgata 711

RESULT 6
AAA79696/c
ID AAA79696 standard; cDNA; 2084 BP.

AC AAA79696;

XX 27-NOV-2000 (first entry)

XX Eucalyptus grandis cell signalling involved polynucleotide SEQ ID NO:848.

XX Eucalyptus grandis; Pinus radiata; Monterey pine; plant; modification;
KW plant cell signalling; modulation; transgenic plant; pathogen; growth;
KW environmental change; development; cell proliferation; differentiation;
KW elongation; survival; disease resistance; nutrient metabolism; ss.

XX Eucalyptus grandis.

OS MO200042171-AL.

PN 20-JUL-2000.

XX 11-JAN-2000; 2000MO-US00724.

PR 12-JAN-1999; 99US-0228986.

PR 01-NOV-1999; 99US-0162866.

PA (GENE-) GENESIS RES & DEV CORP LTD.

XX Strabala TJ, Nieuwenhuizen NJ;

XX

DR WPI; 2000-476052/41.
XX Isolated polynucleotide encoding a polypeptide involved in cell
PT signaling used for generating transgenic plants with modified responses
PT to external signals -
XX
XX Claim 1; Page 408-409; 527pp; English.
XX
CC AAA79263 to AAA79736 and AAB25100 to AAB25570 represent polynucleotide
CC and protein sequences isolated from eucalyptus (Eucalyptus grandis) or
CC pine (Pinus radiata also known as Monterey pine). The protein sequences
CC are involved in cell signaling. The polynucleotide and protein
CC sequences can be used to modify the response of plant cells to external
CC signals e.g. environmental changes or pathogens during the growth and
CC development of a plant. They can be used to modify cell proliferation,
CC differentiation, elongation and survival, resistance to disease and
CC nutrient metabolism. Examples of modifications which can be produced are
CC altered fruit ripening and senescence of leaves and flowers e.g. to
CC delay senescence and prolong the life of cut flowers or enhance
CC senescence of reproductive organs to engineer sterile plants. Other
CC modifications can be used to delay senescence in selected cell types or
CC organs providing fruit and vegetables which have a longer shelf life
CC between harvest and consumption, or to decrease branching frequency in
CC forest tree species giving long stretches of valuable knot-free clear
CC wood which can be used in solid timber furniture and veneers.
XX
SQ Sequence 2084 BP; 502 A; 434 C; 566 G; 582 T; 0 other;

Query Match 80.9%; Score 17.8; DB 21; Length 2084;
Best Local Similarity 90.5%; Pred. No. 19;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 tgttgatcacaaagttagat 21
|||||
DB 1347 tcttgcagcacaagttagat 1327

RESULT 7
AAC18114
ID AAC18114 standard; cDNA; 447 BP.
XX
XX AAC18114;
XX
XX 06-OCT-2000 (first entry)
XX
XX Human secreted protein 5' EST, SEQ ID NO: 22189.
XX
XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX gene therapy; chromosome mapping; ss.
XX
XX Homo sapiens.
XX
XX EP1033401-A2.
XX
XX 06-SEP-2000.
XX
XX 21-FEB-2000; 2000EP-0200610.
XX
XX 26-FEB-1999; 990US-0122487.
XX
XX (GENSET) GENSET.
XX
XX Dumas Milne Edwards J, Duclert A, Giordano J;
XX
XX WPI; 2000-500381/45.
XX
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
XX diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
XX Claim 1; SEQ ID 22189; 71pp + CD-ROM; English.
XX

CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
XX expression and secretion vectors.
XX
SQ Sequence 447 BP; 187 A; 57 C; 66 G; 128 T; 9 other;

Query Match 78.2%; Score 17.2; DB 21; Length 447;
Best Local Similarity 86.4%; Pred. No. 32;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 tgttgatcacaaagttagata 22
|||||
DB 240 tgttgatcacaaacttaata 261

RESULT 8
AAC42365/C
ID AAC42365 standard; DNA; 1290 BP.
XX
XX AAC42365;
XX
XX 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 35275.
XX
XX Hybridisation assay; genetic mapping; gene expression control;
XX protein identification; signal transduction pathway;
XX metabolic pathway; promoter; termination sequence; ss.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 990US-0121825.
XX 05-MAR-1999; 990US-0123180.
XX 09-MAR-1999; 990US-0123548.
XX 23-MAR-1999; 990US-0125788.
XX 25-MAR-1999; 990US-0126264.
XX 29-MAR-1999; 990US-0126785.
XX 01-APR-1999; 990US-0127462.
XX 06-APR-1999; 990US-0128234.
XX 08-APR-1999; 990US-0128714.
XX 16-APR-1999; 990US-0129845.
XX 19-APR-1999; 990US-0130077.
XX 21-APR-1999; 990US-0130449.
XX 23-APR-1999; 990US-0130510.
XX 23-APR-1999; 990US-0130891.
XX 28-APR-1999; 990US-0131449.
XX 30-APR-1999; 990US-0132048.
XX 30-APR-1999; 990US-0132407.
XX 04-MAY-1999; 990US-0132484.
XX 05-MAY-1999; 990US-0132485.
XX 06-MAY-1999; 990US-0132486.
XX 07-MAY-1999; 990US-0132487.
XX 11-MAY-1999; 990US-0132863.
XX 14-MAY-1999; 990US-0134256.
XX

PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137724.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.

PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149929.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149920.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150566.
PR 25-AUG-1999; 99US-0150884.
PR 26-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.

PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149358.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151085.
 PR 27-AUG-1999; 99US-0151086.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 01-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 22-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.

PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.
 Query Match 78.2%; Score 17.2; DB 21; Length 1413;
 Best Local Similarity 86.4%; Pred. No. 36;
 Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 ttttgatcacaaagatttgata 22
 Db 576 TGTGTACACAGATTGTGA 555
 RESULT 10
 AAS22799 standard; CDNA; 2038 BP.
 ID AAS22799;
 AC AAS22799;
 XX
 XX
 DT 24-OCT-2001 (first entry)
 XX
 DE Human cDNA encoding a novel human protein #365.
 XX
 KW Human; novel protein; se; Antianemic; osteoparhic; antiinflammatory;
 KW immunomodulatory; cytosolic; neuroprotective; vulnery; neotropic;
 KW anticonvulsant; antiallergic; cerebroprotective; antifungal; antiviral;
 KW antibacterial; antiallergic; dermatological; hemostatic; antisthmatic;
 KW thrombolytic; immunogen; antibody; gene therapy; neurological disorder;
 KW Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;
 KW tissue regeneration; immune disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200155437-A2.
 XX
 PD 02-AUG-2001.
 XX
 PF 25-JAN-2001; 2001MO-US02623.
 XX
 XX 25-JAN-2000; 2000US-0491404.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 XX
 DR WPI; 2001-451939/48.
 XX
 PT P-PSDB; AAU14494.
 XX
 PT Isolated polypeptides useful for treating anti-inflammatory diseases,
 XX nervous system disorders, and for regenerating bone and cartilage -
 PS Claim 1; Page 733-734; 894pp; English.
 XX
 CC The invention relates to polynucleotides encoding novel human
 CC proteins or their active domains. The polypeptides, polynucleotides and
 CC antibodies raised against the polypeptides are used in a method of
 CC treatment of a mammal and prevention of disorders caused by the aberrant
 CC protein expression or activity. The polypeptides can be used as
 CC molecular weight markers, food supplements, and in antibody production.
 CC The polypeptides are used to identify compounds which bind to the
 CC polypeptides. Polynucleotides of the invention are used as probes and
 CC primers, for sequencing, for chromosome or gene mapping, in the
 CC production of recombinant proteins, and in generating anti-sense DNA or
 CC RNA and in gene therapy. Polypeptides of the invention can be used to
 CC target drugs to a tumour, in assays to determine biological activity, to
 CC raise antibodies/elicit an immune response, to determine receptors or ligands.
 CC protein levels, as tissue markers, and to isolate receptors or ligands.
 CC Polypeptides of the invention may also be useful in treating platelet
 CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,
 CC ligament and/or nerve tissue, wound healing, treating burns, promoting
 CC the proliferation, differentiation and survival of stem cells, as a
 CC contraceptive, treating osteoporosis and osteoarthritis, anaemia,
 CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral

CC scleriosis, stroke, immune deficiencies resulting from bacterial, viral or
CC fungal infection or from autoimmunity, cancer, allergy, asthma,
CC graft-versus-host disease, eczema, haemophilia, thrombosis,
CC anti-inflammatory diseases, nervous system disorders, and infection.
CC The present sequence encodes a protein of the invention.

XX
SQ Sequence 2038 BP; 658 A; 400 C; 503 G; 477 T; 0 other;

Query Match 78.2%; Score 17.2; DB 22; Length 2038;

Best Local Similarity 86.4%; Pred. No. 38;

Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 ttttgatcacagaattgata 22
||||| |||||||

DB 1929 ttttgatcacagaattttaa 1950

RESULT 11

AAH15352/C
ID AAH15352 standard; CDNA; 1764 BP.

AC AAH15352;

DT 26-JUN-2001 (first entry)

DE Human cDNA sequence SEQ ID NO:13526.

XX Human; primer: detection; diagnosis; antisense therapy; gene therapy; ss.

KW Homo sapiens.

XX EP1074617-A2.

PD 07-FEB-2001.

PF 28-JUL-2000; 2000EP-0116126.

XX 29-JUL-1999; 99JP-0248036.

PR 27-AUG-1999; 98JP-0300253.

PR 11-JAN-2000; 2000JP-0118776.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

PA (HELI-) HELIX RES INST.

PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

DR MPI; 2001-318749/34.

PT Primer sets for synthesizing polynucleotides, particularly the 5602

PT full-length cDNAs defined in the specification, and for the detection

PT and/or diagnosis of the abnormality of the proteins encoded by the

PT full-length cDNAs -

CC Claim 8; SEQ ID 13526; 2537bp + CD ROM; English.

CC The present invention describes primer sets for synthesizing 5602

CC full-length cDNAs defined in the specification. Where a primer set

CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary

CC to the complementary strand of a polynucleotide which comprises one of

CC the 5602 nucleotide sequences defined in the specification, where the

CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination

CC of an oligonucleotide comprising a sequence complementary to the

CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.

XX
SQ Sequence 1764 BP; 511 A; 347 C; 392 G; 514 T; 0 other;

Query Match 76.4%; Score 16.8; DB 22; Length 1764;

Best Local Similarity 90.0%; Pred. No. 58;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ttttgatcacagaattgata 20
||||| |||||||

DB 783 TGTGTGATGACCGAGATTGA 764

RESULT 12

AAT90543/C
ID AAT90543 standard; DNA; 2582 BP.

AC AAT90543;

DT 13-FEB-1998 (first entry)

DE Maize male fertility gene Z33 Zm41-A.

XX Zm41-A; male fertility; hybrid seed; transgenic plant; antisense;

KW ribozyme; male sterile; maize; Ms41-A; ds.

XX Zea mays.

OS

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Location/Qualifiers

Key

FT exon

FT intron

FT exon

FT intron

FT exon

FT intron

FT exon

FT intron

FT exon

FT intron

FT exon

FT intron

FT exon

FT intron

FT exon

FT intron

FT exon

FT intron

FT exon

FT intron

FT exon

WO9723618-A1.

PD 03-JUL-1997.

PF 20-DEC-1996; 96WO-GB03191.

PR 21-DEC-1995; 95GB-0026218.

PA (GENE-) GENE SHEARS PTY LTD.

PI Baudot G, Garcia D, Hodge R, Perez P;

PI MPI; 1997-351055/32.

PT Nucleic acid encoding proteins involved in male fertility in plants

PT - used to control fertility and for production of hybrid seed

PT Claim 3; Fig 15; 85bp; English.

CC This DNA sequence comprises the maize Z33 Zm41-A gene. This is an

CC orthologue of Arabidopsis Ms41-A (see AAT90522), a gene that confers

CC male fertility. It was obtained following a database search for

CC sequences that showed homology to Ms41-A DNA, isolation of a

CC partial cDNA clone, and use of this clone to isolate Zm41-A genes

CC Z31 (AAT90542), Z33 (AAT90543) and Z35 (AAT90544) from genomic lambda

CC libraries. The genes show a high level of conservation. Z35 may

be derived from 231 via genetic rearrangements, deletions and/or insertions. 233 has subsequent deletions from 235 and is truncated, having only exons 3, 5 and 6. The M41-A and Zm41-A genes, antisense or ribozyme sequences can be used to produce transgenic plants with controlled male fertility. Male sterile plants are useful for hybrid seed production, particularly in plants where restoration of fertility is not needed, e.g. Brassicaceae, lettuce, spinach and onions.

Sequence 2582 BP; 670 A; 512 C; 537 G; 863 T; 0 other;

Query Match 76.4%; Score 16.8; DB 18; Length 2582;
Best Local Similarity 90.0%; Pred. No. 60;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 tgttgatcacagaattga 20
|||||
Db 268 TGTGATCACAGATCTGA 249

RESULT 13

AAC56059 standard; DNA; 2444 BP.
AAC56059;

25-JAN-2001 (first entry)

Eucalyptus grandis transcription factor DNA sequence #190.

Plant; transcription factor; gene expression; eucalyptus; pine; acacia; poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor; basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS; homeodomain zipper; LIM domain; AP2; EREBs; zinc finger domain; type 2 Cys2His2; CCAAT box element; MYB; ss.

Eucalyptus grandis.

WO200053724-A2.

14-SEP-2000.
09-MAR-2000; 2000MO-US06112.
11-MAR-1999; 99US-0266513.
18-AUG-1999; 99US-0149485.

(GENE-) GENESIS RES & DEV CORP LTD.
(FLET-) FLETCHER CHALLENGE FORESTS LTD.

Wood M, McGrath A, Shenk MA, Glenn M;

WPI; 2000-579369/54.

New isolated polynucleotide encoding a plant transcription factor for producing a plant e.g. a woody plant, preferably eucalyptus or pine, having modified gene expression or modified activity of a polypeptide

Claim 1; Pages 94-95; 747pp; English.

The present invention relates to novel plant transcription factors from Eucalyptus grandis or Pinus radiata. The present sequence is the coding sequence for one such transcription factor. The transcription factor may be used to produce a plant having modified gene expression such as a woody plant e.g. a eucalyptus, pine, acacia, poplar, sweetgum, teak, or mahogany species or to modify the activity of a polypeptide in a plant. The transcription factors of the present invention are members from the following families of regulatory proteins: bZIP, bZIP family of G-box binding factors, basic helix-loop-helix zipper, LIM domain, AP2 homeotic/homeodomain/homeobox/MADS, homeodomain zipper, LIM domain, AP2 and EREBs, zinc finger domains of type 2 Cys2His2, CCAAT box elements

CC and MYB.
XX Sequence 2444 BP; 643 A; 554 C; 640 G; 607 T; 0 other;
SQ

Query Match 74.5%; Score 16.4; DB 21; Length 2444;
Best Local Similarity 94.4%; Pred. No. 94;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 tgttgatcacagaattc 18
|||||
Db 1451 tgaatgacacagaattc 1468

RESULT 14
AAH3702/c
ID AAH3702 standard; CDNA; 335 BP.

AAH3702;

03-SEP-2001 (first entry)

Human colon cancer antigen encoding cDNA SEQ ID NO:758.

Human; colon cancer; colon cancer antigen; diagnosis; detection; colorectal carcinoma; ss.

Homo sapiens.

WO200122920-A2.

05-APR-2001.

28-SEP-2000; 2000MO-US26524.

29-SEP-1999; 99US-0157137.

03-NOV-1999; 99US-0163280.

(HUMA-) HUMAN GENOME SCI INC.

Ruben SM, Barash SC, Birse CE, Rosen CA;

WPI; 2001-235357/24.

Nucleic acids encoding 427 human colon cancer-associated polypeptides, useful for preventing, diagnosing and/or treating colorectal cancers - Claim 1; Page 2738; 9803pp; English.

AAH32943 to AAH37195 and AAG77788 represent human colon cancer-associated nucleic acid molecules (N) and proteins (P), where the proteins are collectively known as colon cancer antigens. The colon cancer antigens have cytostatic activity and can be used in gene therapy and vaccine production. N and P may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate P expression. For example, N and P may be used to treat disorders CC associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of P by expressing in inactive proteins or to supplement the patients own production of P. Additionally, N may be used to produce the colon cancer-associated P, by inserting the nucleic acids into a host cell and culturing the cell to express the proteins. N and P can be used in the prevention, diagnosis and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204 and AAG77789 represent sequences used in the exemplification of the present invention.

N.B. Pages 666 to 682 and page 7053 of the sequence listing were missing at time of publication, meaning no sequences are present for SEQ ID NO:1027 to 1052, 7921 and 7922.

Sequence 335 BP; 95 A; 71 C; 60 G; 105 T; 4 other;

CC restenosis, prostate diseases, obesity, disorders involving osteoclasts
 CC such as osteoporosis, arthritis or malignancies, diseases of testes,
 CC lung or thymus, digestive/endocrine disorders, infections and AIDS. The
 CC polypeptides are also useful for identifying their binding partners.
 CC The present sequence represents a gene encoding a human secreted protein
 CC (see descriptor line for gene number and clone identification).

XX
 SQ Sequence 951 BP; 296 A; 134 C; 192 G; 329 T; 0 other;

Query Match 73.6%; Score 16.2; DB 20; Length 951;
 Best Local Similarity 85.7%; Pred. No. 1.1e+02;

Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 tgltgatcacaaagttagat 21
 ||||| || ||||| |||||
 Db 524 TGTGGTTCCTCAAGATCTGAT 504

Search completed: December 15, 2001, 03:29:47
 Job time: 7887 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 03:22:49 ; Search time 172.39 Seconds
(without alignments)
28,903 Million cell updates/sec

Title: US-09-380-826A-7
Perfect score: 22
Sequence: 1 ttttgatcaccaagattgata 22

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 11323899 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2-6/ptodata/2/1na/5A_COMB.seq:*
2: /cgn2-6/ptodata/2/1na/5B_COMB.seq:*
3: /cgn2-6/ptodata/2/1na/6A_COMB.seq:*
4: /cgn2-6/ptodata/2/1na/6B_COMB.seq:*
5: /cgn2-6/ptodata/2/1na/PCUTUS_COMB.seq:*
6: /cgn2-6/ptodata/2/1na/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB* ID	Description
c 1	16.8	76.4	2582	4 US-09-102-528-31	Sequence 31, Appl
c 2	16.2	73.6	2000	1 US-08-014-326-1	Sequence 1, Appl
c 3	16.2	73.6	2001	3 US-08-417-089-1	Sequence 1, Appl
c 4	16.2	73.6	2001	3 US-08-695-651-1	Sequence 1, Appl
c 5	16.2	73.6	2001	4 US-08-930-285-1	Sequence 1, Appl
c 6	16.2	73.6	2001	4 US-08-695-421-1	Sequence 1, Appl
c 7	16.2	73.6	4345	2 US-08-244-537-1	Sequence 1, Appl
c 8	16.2	73.6	7470	3 US-08-417-089-5	Sequence 1, Appl
c 9	16.2	73.6	7470	3 US-08-695-651-5	Sequence 5, Appl
c 10	16.2	73.6	7470	4 US-08-930-285-5	Sequence 5, Appl
c 11	16.2	73.6	7470	4 US-08-695-421-5	Sequence 5, Appl
c 12	15.6	70.9	296	3 US-08-602-145-14	Sequence 14, Appl
c 13	15.6	70.9	1095	4 US-08-928-383B-3	Sequence 3, Appl
c 14	15.6	70.9	1096	1 US-08-684-862-8	Sequence 8, Appl
c 15	15.6	70.9	1584	4 US-08-928-383B-1	Sequence 1, Appl
c 16	15.6	70.9	1989	2 US-08-792-055-1	Sequence 1, Appl
c 17	15.6	70.9	2434	4 US-09-272-486-1	Sequence 1, Appl
c 18	15.6	70.9	2770	4 US-09-008-697A-13	Sequence 13, Appl
c 19	15.6	70.9	7739	4 US-09-195-966-1	Sequence 1, Appl
c 20	15.6	70.9	7745	1 US-08-299-675-1	Sequence 1, Appl
c 21	15.6	70.9	7745	1 US-08-485-241-1	Sequence 1, Appl
c 22	15.6	70.9	7745	2 US-08-874-162-1	Sequence 1, Appl
c 23	15.2	69.1	1016	1 US-08-399-986B-3	Sequence 3, Appl
c 24	15.2	69.1	1016	1 US-08-493-754A-3	Sequence 3, Appl
c 25	15.2	69.1	2182	1 US-08-399-986B-1	Sequence 1, Appl
c 26	15.2	69.1	2182	1 US-08-493-754A-1	Sequence 1, Appl
c 27	15.2	69.1	2598	3 US-08-745-892-20	Sequence 20, Appl

c 28	15.2	69.1	5420	6 5256642-3	Patent No. 5256642
c 29	15.2	69.1	5420	6 5472939-3	Patent No. 5472939
c 30	15.2	69.1	5872	4 US-09-102-528-32	Sequence 32, Appl
c 31	15.2	69.1	6951	6 5256642-1	Patent No. 5256642
c 32	15.2	69.1	6951	6 5472939-1	Patent No. 5472939
c 33	14.8	67.3	337	4 US-08-991-789A-13	Sequence 13, Appl
c 34	14.8	67.3	1855	3 US-08-961-083-71	Sequence 71, Appl
c 35	14.8	67.3	3159	1 US-08-119-361-4	Sequence 4, Appl
c 36	14.8	67.3	3159	3 US-08-336-308A-3	Sequence 3, Appl
c 37	14.8	67.3	3159	3 US-08-822-324-3	Sequence 3, Appl
c 38	14.8	67.3	3159	4 US-09-490-931-3	Sequence 3, Appl
c 39	14.8	67.3	7266	3 US-08-336-308A-9	Sequence 9, Appl
c 40	14.8	67.3	7266	3 US-08-822-324-5	Sequence 5, Appl
c 41	14.8	67.3	7266	4 US-09-490-931-9	Sequence 9, Appl
c 42	14.8	67.3	8640	1 US-08-570-311-28	Sequence 28, Appl
c 43	14.8	67.3	9919	3 US-08-880-179-1	Sequence 1, Appl
c 44	14.6	66.4	293	3 US-08-866-340-13	Sequence 13, Appl
c 45	14.6	66.4	293	4 US-09-103-875-17	Sequence 17, Appl

ALIGNMENTS

```
RESULT 1
US-09-102-528-31/c
; Sequence 31, Application US/09102528
; Patent No. 6207883
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: DNA Sequences
; NUMBER OF SEQUENCES: 32
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30 (EPC)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/102,528
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/GB96/03191
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2582 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-102-528-31

Query Match 76.4%; Score 16.8; DB 4; Length 2582;
Best Local Similarity 90.0%; Pred. No. 14;
Matches 18; Conservative 2; Indels 0; Gaps 0;

QY 1 ttttgatcaccaagattga 20
Db 268 ttttgatcaccaagattga 249

RESULT 2
US-08-014-326-1
; Sequence 1, Application US/08014326
; Patent No. 5498544
; GENERAL INFORMATION:
; APPLICANT: Gengenbach, Burle G.
; APPLICANT: Somers, David A.
; APPLICANT: Wyse, Donald L.
; APPLICANT: Gronwald, John W.
; APPLICANT: Egli, Margaret A.
; APPLICANT: Lutz, Shiela M.
; TITLE OF INVENTION: Method and An Acetyl CoA Carboxylase
```

```
; TITLE OF INVENTION: Alteration in Oil Content of Plants
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESSES:
; ADDRESS: Merchant & Gould
; STREET: 3100 No. 5498544west Center.
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/014,326
; FILING DATE: 05-FEB-1993
; CLASSIFICATION: 800
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 07/917,462
; FILING DATE: 21-JUL-1992
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 07/538,674
; FILING DATE: 18-JUN-1990
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 07/269,584
; FILING DATE: 10-NOV-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Kowalchuk, Katherine M.
; REGISTRATION NUMBER: 36,848
; REFERENCE/DOCKET NUMBER: 600,258-US01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-332-5300
; TELEFAX: 612-332-9081
; INFORMATION FOR SEQ. ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2000 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: 2 kb fragment of lambda clone #15-14
;
US-08-014-326-1
;
Query Match 73.6%; Score 16.2; DB 1; Length 2000;
Best Local Similarity 85.7%; Pred. No. 27;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
;
QY 2 gtcgacacagaattgata 22
;
Db 42 GTTGATGACACAGAGTGTGA 62
;
;
RESULT 3
US-08-417-089-1
; Sequence 1, Application US/08417089
; Patent No. 6069298
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: METHODS AND AN ACETYL COA CARBOXYLASE GENE
; TITLE OF INVENTION: FOR CONFERRING HERBICIDE TOLERANCE AND AN ALTERATION IN
; NUMBER OF SEQUENCES: 11
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (RPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/417,089
; INFORMATION FOR SEQ. ID NO: 1:
```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2001 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
;
US-08-417-089-1
;
Query Match 73.6%; Score 16.2; DB 3; Length 2001;
Best Local Similarity 85.7%; Pred. No. 27;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
;
QY 2 gtcgacacagaattgata 22
;
Db 42 GTTGATGACACAGAGTGTGA 62
;
;
RESULT 4
US-08-695-651-1
; Sequence 1, Application US/08695651
; Patent No. 6146867
; GENERAL INFORMATION:
; APPLICANT: Gengenbach, B. G.
; APPLICANT: Somers, D. A.
; APPLICANT: Wyse, D. L.
; APPLICANT: Gronwald, J. W.
; APPLICANT: Egli, M. A.
; APPLICANT: Lutz, S. M.
; TITLE OF INVENTION: METHOD FOR PRODUCING PLANT ACETYL COA
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P. A.
; STREET: P.O. Box 2938.
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/695,651
; FILING DATE: 12-AUG-1996
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 08/417089
; FILING DATE: 05-APR-1995
; APPLICATION NUMBER: 08/014326
; FILING DATE: 05-FEB-1993
; APPLICATION NUMBER: 07/917462
; FILING DATE: 21-JUL-1992
; APPLICATION NUMBER: 07/538674
; FILING DATE: 18-JUN-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Woessner, Warren D
; REGISTRATION NUMBER: 30,440
; REFERENCE/DOCKET NUMBER: 600,318US3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-373-6900
; TELEFAX: 612-339-3061
; TELEX:
; INFORMATION FOR SEQ. ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2001 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
;
US-08-695-651-1
```

Query Match 73.6%; Score 16.2; DB 3; Length 2001;
Best Local Similarity 85.7%; Pred. No. 27;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 gttgatacacaagatttgata 22
||||| ||||| ||| ||
DB 42 GTTGATGACACAGACTTGTTA 62

RESULT 5
US-08-930-285-1
Sequence 1, Application US/08930285
Patent No. 6222099

GENERAL INFORMATION:

APPLICANT: Regents of the University of Minnesota, et al.
TITLE OF INVENTION: TRANSGENIC PLANTS EXPRESSING ACETYL COA CARBOXYLASE GE

NUMBER OF SEQUENCES: 26

CORRESPONDENCE ADDRESS:

ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P. A.

STREET: P. O. Box 2938

CITY: Minneapolis

STATE: MN

COUNTRY: USA

ZIP: 55402

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/930.285

FILING DATE: 13-APR-1998

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US96/04625

FILING DATE: 04-APR-1996

ATTORNEY/AGENT INFORMATION:

NAME: Embretson, Janet E.

REGISTRATION NUMBER: 39,665

REFERENCE/DOCKET NUMBER: 600.318054

TELECOMMUNICATION INFORMATION:

TELEPHONE: 612-339-0331

TELEFAX: 612-339-3061

TELEX:

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 2001 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

HYPOTHEICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE:

ORIGINAL SOURCE:

US-08-930-285-1

Query Match 73.6%; Score 16.2; DB 4; Length 2001;

Best Local Similarity 85.7%; Pred. No. 27;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 gttgatacacaagatttgata 22
||||| ||||| ||| ||
DB 42 GTTGATGACACAGACTTGTTA 62

RESULT 6

US-08-695-421-1
Sequence 1, Application US/08695421
Patent No. 6268550

GENERAL INFORMATION:

APPLICANT: Gengenbach, B. G.
APPLICANT: Somers, D. A.
APPLICANT: Wyse, D. L.
APPLICANT: Gronwald, J. W.
APPLICANT: Egli, M. A.
APPLICANT: Lutz, S. M.

TITLE OF INVENTION: METHODS AND AN ACETYL COA CARBOXYLASE GENE
FOR CONFERRING HERBICIDE TOLERANCE AND AN ALTERATION IN OIL

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P. A.

STREET: P. O. Box 2938

CITY: Minneapolis

STATE: MN

COUNTRY: USA

ZIP: 55402

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/695.421

FILING DATE: 23-AUG-1996

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/417089

FILING DATE: 05-APR-1995

APPLICATION NUMBER: 08/014326

FILING DATE: 05-FEB-1993

APPLICATION NUMBER: 07/917462

FILING DATE: 21-JUL-1992

APPLICATION NUMBER: 07/538674

FILING DATE: 18-JUN-1990

ATTORNEY/AGENT INFORMATION:

NAME: Woessner, Warren D

REGISTRATION NUMBER: 30,440

REFERENCE/DOCKET NUMBER: 600.318052

TELECOMMUNICATION INFORMATION:

TELEPHONE: 612-373-6900

TELEFAX: 612-339-3061

TELEX:

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 2001 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

US-08-695-421-1

Query Match 73.6%; Score 16.2; DB 4; Length 2001;

Best Local Similarity 85.7%; Pred. No. 27;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 gttgatacacaagatttgata 22
||||| ||||| ||| ||
DB 42 GTTGATGACACAGACTTGTTA 62

RESULT 7

US-08-244-537-1
Sequence 1, Application US/08244537
Patent No. 5854420

GENERAL INFORMATION:

APPLICANT: ASHTON, ANTHONY R.

APPLICANT: JENKINS, COLIN L.D.

APPLICANT: WHITEFIELD, PAUL R.

TITLE OF INVENTION: MAIZE ACETYL COA CARBOXYLASE ENCODING

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P.
STREET: 1100 New York Avenue, N.W.
CITY: Washington
STATE: D. C.
COUNTRY: U.S.A.
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/244,537
FILING DATE: 18-AUG-1994
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9125330.2
FILING DATE: 28-NOV-1991
PRIOR APPLICATION DATA: PCT/GB92/02205
FILING DATE: 27-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: KOKULIS, PAUL N.
REGISTRATION NUMBER: 16,773
REFERENCE/DOCKET NUMBER: 203094/SEE 36663/UST
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4345 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-244-537-1

Query Match 73.6%; Score 16.2; DB 2; Length 4345;
Best Local Similarity 85.7%; Pred. No. 30;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 gttgatacacaagatttgata 22
||||| ||||| ||| ||

Db 902 GTTGATGACAGAAGTTGTTA 922

RESULT 8
US-08-417-089-5
Sequence 5, Application US/08417089
Patent No. 6069298
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: METHODS AND AN ACETYL COA CARBOXYLASE GENE
TITLE OF INVENTION: FOR CONFERRING HERBICIDE TOLERANCE AND AN ALTERATION IN
TITLE OF INVENTION: OIL CONTENT OF PLANTS
NUMBER OF SEQUENCES: 11
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/417,089
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 7470 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-417-089-5

Query Match 73.6%; Score 16.2; DB 3; Length 7470;
Best Local Similarity 85.7%; Pred. No. 32;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 gttgatacacaagatttgata 22
||||| ||||| ||| ||

Db 3972 GTTGATGACAGAAGTTGTTA 3992

RESULT 9
US-08-695-651-5
Sequence 5, Application US/08695651
Patent No. 6146867
GENERAL INFORMATION:
APPLICANT: Gengenbach, B. G.
APPLICANT: Somers, D. A.
APPLICANT: Wyse, D. L.
APPLICANT: Gronwald, J. W.
APPLICANT: Egli, M. A.
APPLICANT: Lutz, S. M.
TITLE OF INVENTION: METHOD FOR PRODUCING PLANT ACETYL COA
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
STREET: P.O. Box 2938
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55402
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/695,651
FILING DATE: 12-AUG-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/417089
FILING DATE: 05-APR-1995
APPLICATION NUMBER: 08/014326
FILING DATE: 05-FEB-1993
APPLICATION NUMBER: 07/917462
FILING DATE: 21-JUL-1992
APPLICATION NUMBER: 07/538674
FILING DATE: 18-JUN-1990
ATTORNEY/AGENT INFORMATION:
NAME: Woessner, Warren D
REGISTRATION NUMBER: 30,440
REFERENCE/DOCKET NUMBER: 600.318053
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-373-6900
TELEFAX: 612-339-3061
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 7470 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-695-651-5

Query Match: 73.6%; Score 16.2; DB 3; Length 7470;
Best Local Similarity 85.7%; Pred. No. 32;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 gttgatacacaagatttgata 22
||||| ||||| ||| ||

DB 3972 GTTGATGACAGAGTTGTTA 3992

RESULT 10

US-08-930-285-5

; Sequence 5, Application US/08930285

; Patent No. 6222099

; GENERAL INFORMATION:

; APPLICANT: Regents of the University of Minnesota, et al.

; TITLE OF INVENTION: TRANSGENIC PLANTS EXPRESSING ACETYL COA CARBOXYLASE GE

; NUMBER OF SEQUENCES: 26

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P. A.

; STREET: P. O. Box 2938

; CITY: Minneapolis

; STATE: MN

; COUNTRY: USA

; ZIP: 55402

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FASTSEQ Version 1.5

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/930,285

; FILING DATE: 13-APR-1998

; CLASSIFICATION: 800

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: PCT/US96/04625

; FILING DATE: 04-APR-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Embretson, Janet E.

; REGISTRATION NUMBER: 39,665

; REFERENCE/DOCKET NUMBER: 600.318054

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 612-339-0331

; TELEFAX: 612-339-3061

; TELEX:

; INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 7470 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; HYPOETHERICAL: NO

; ANTI-SENSE: NO

; FRAGMENT TYPE:

; ORIGINAL SOURCE:

US-08-930-285-5

Query Match

Best Local Similarity

Matches 18; Conservative

OY 2 gtggatcacaaagtgtgata 22

DB 3972 GTTGATGACAGAGTTGTTA 3992

RESULT 11

US-08-695-421-5

; Sequence 5, Application US/08695421

; Patent No. 6268550

; GENERAL INFORMATION:

; APPLICANT: Gengenbach, B. G.

; APPLICANT: Somers, D. A.

; APPLICANT: Wyse, D. L.

; APPLICANT: Gronwald, J. W.

; APPLICANT: Egli, M. A.

; APPLICANT: Lutz, S. M.

; TITLE OF INVENTION: METHODS AND AN ACETYL COA CARBOXYLASE GENE

; TITLE OF INVENTION: FOR CONFERRING HERBICIDE TOLERANCE AND AN ALTERATION IN OIL

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.

; STREET: P.O. Box 2938

; CITY: Minneapolis

; STATE: MN

; COUNTRY: USA

; ZIP: 55402

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FASTSEQ Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/695,421

; FILING DATE: 23-AUG-1996

; CLASSIFICATION: 800

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: 08/417089

; FILING DATE: 05-APR-1995

; APPLICATION NUMBER: 08/014326

; FILING DATE: 05-FEB-1993

; APPLICATION NUMBER: 07/917462

; FILING DATE: 21-JUL-1992

; APPLICATION NUMBER: 07/538674

; FILING DATE: 18-JUN-1990

; ATTORNEY/AGENT INFORMATION:

; NAME: Woessner, Warren D

; REGISTRATION NUMBER: 30,440

; REFERENCE/DOCKET NUMBER: 600.318052

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 612-373-6900

; TELEFAX: 612-339-3061

; TELEX:

; INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 7470 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

US-08-695-421-5

Query Match

Best Local Similarity

Matches 18; Conservative

OY 2 gtggatcacaaagtgtgata 22

DB 3972 GTTGATGACAGAGTTGTTA 3992

RESULT 12

US-08-602-145-14/c

; Sequence 14, Application US/08602145

; Patent No. 6025336

; GENERAL INFORMATION:

; APPLICANT: Goltzy, Kristin L.

; APPLICANT: Greenberger, Joel S.

; TITLE OF INVENTION: DETERMINING EXPOSURE TO IONIZING RADIATION

; NUMBER OF SEQUENCES: 14

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Foley & Lardner

; STREET: 3000 K Street, N.W., Suite 500

; CITY: Washington

; STATE: D.C.

; COUNTRY: USA

; ZIP: 20007-5109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

Oy 1 tgttgatcacagaatttgata 22
 ||| | ||||| |||||
 Db 112 TGTAGTATCACAAGGTTTCCTA 91

RESULT 15

US-08-928-383B-1/c
 ; Sequence 1, Application US/08928383B
 ; Patent No. 6210921
 ; GENERAL INFORMATION:
 ; APPLICANT: Robert W. Finberg, Jeffrey M. Bergelson,
 ; APPLICANT: and Marshall S. Horwitz
 ; TITLE OF INVENTION: CAR, A No. 6210921el Cocksacklevirus and Adenovirus
 ; TITLE OF INVENTION: Receptor
 ; NUMBER OF SEQUENCES: 26
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: LAHIVE & COCKFIELD, LLP
 ; STREET: 28 State Street
 ; CITY: Boston
 ; STATE: Massachusetts
 ; COUNTRY: USA
 ; ZIP: 02109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/928,383B
 ; FILING DATE: 12-SEP-1997
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 60/026,100
 ; FILING DATE: 13-SEP-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Mandragouras, Amy E.
 ; REGISTRATION NUMBER: 36,207
 ; REFERENCE/DOCKET NUMBER: DFN-020
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617)227-7400
 ; TELEFAX: (617)742-4214
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1584 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 60..1157
 ; US-08-928-383B-1

Query Match 70.9%; Score 15.6; DB 4; Length 1584;
 Best Local Similarity 81.8%; Pred. No. 51;
 Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Oy 1 tgttgatcacagaatttgata 22
 ||| | ||||| |||||
 Db 371 TGTATGATCACCAGATTGTGAGA 350

Search completed: December 15, 2001, 03:22:50
 Job time: 7565 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 02:33:48 ; Search time 4441.54 Seconds
(without alignments)
53.226 Million cell updates/sec

Title: US-09-380-826a-7
Perfect score: 22
Sequence: 1 tftgtatcaccaagatttgata 22

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues
Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estopl:*
6: em_estlba:*
7: em_estro:*
8: em_estov:*
9: em_hcc:*
10: qb_est1:*
11: qb_est2:*
12: qb_hcc:*
13: qb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_dln:*
18: em_gss_pro:*
19: em_gss_fod:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	18.8	85.5	361	10	AV532040 AV532040
c 2	18.8	85.5	394	11	226559 APTS1651 Ve
c 3	18.8	85.5	490	10	AB038725 AB038725
c 4	18.8	85.5	527	10	AV520789 AV520789
c 5	18.8	85.5	530	10	A1995637 A1995637
c 6	18.4	83.6	608	11	BE977618 bs66h08.y
c 7	18.4	83.6	934	13	CNS02BWA
c 8	17.8	80.9	600	11	BF006444 BF006444
c 9	17.8	80.9	633	11	BF006254 EST434752
c 10	17.4	79.1	286	10	BB288727 BB288727
c 11	17.4	79.1	407	10	BE581720 BE581720
c 12	17.4	79.1	529	13	AQ497779 HS_5066_B

c 13	17.4	79.1	646	10	AW695022
c 14	17.4	79.1	671	11	BG44853
c 15	17.4	79.1	762	10	BE283043
c 16	17.4	79.1	837	10	BE642809
c 17	17.4	79.1	1101	13	CNS00D26
c 18	17.2	78.2	151	10	BE068225
c 19	17.2	78.2	425	10	A104911
c 20	17.2	78.2	431	13	A0595814
c 21	17.2	78.2	432	13	A0221715
c 22	17.2	78.2	437	11	W43212
c 23	17.2	78.2	439	13	A0796989
c 24	17.2	78.2	500	10	A1733664
c 25	17.2	78.2	524	13	B62582
c 26	17.2	78.2	531	13	A0223477
c 27	17.2	78.2	536	10	AM034253
c 28	17.2	78.2	546	10	A1779714
c 29	17.2	78.2	579	13	B67515
c 30	17.2	78.2	602	10	A1777095
c 31	17.2	78.2	614	10	BE187570
c 32	17.2	78.2	619	10	AM963686
c 33	17.2	78.2	638	11	BE346432
c 34	17.2	78.2	680	13	B57784
c 35	17.2	78.2	698	13	A2193715
c 36	17.2	78.2	773	13	CNS07DUS
c 37	17.2	78.2	888	13	CNS079F4
c 38	17.2	78.2	904	13	CNS029DN
c 39	17.2	78.2	949	13	CNS04HNP
c 40	17.2	78.2	949	13	CNS04HNP
c 41	17.2	78.2	949	13	CNS04HNP
c 42	16.8	76.4	238	10	BE118999
c 43	16.8	76.4	260	11	T20362
c 44	16.8	76.4	296	11	T20362
c 45	16.8	76.4	308	11	BE398487

ALIGNMENTS

RESULT 1
AV532040/c
LOCUS
DEFINITION
thaliana CDNA clone FB034f09f 3', mRNA sequence.
ACCESSION
AV532040
VERSION
AV532040.1 GI:8692323
KEYWORDS
EST
SOURCE
thale cress.
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
1 (bases 1 to 361)
Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries
JOURNAL
DNA Res. 7, 175-180 (2000)
MEDLINE
20363093
COMMENT
Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
FEATURES
source
1..361
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="FB034f09f"
/clone_lib="Arabidopsis thaliana flower buds Columbia"
/tissue_type="flower buds"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT 157 a 111 c 101 g 158 t
 ORIGIN

Query Match 85.5%; Score 18.8; DB 10; Length 527;
 Best Local Similarity 90.9%; Pred. No. 1.1e+02;
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ttttgatcacaaagtctgata 22
 ||||| ||||| ||||| |||||
 Db 247 TGTGTGTTCCACAAGATTGACA 226

RESULT 5
 A1995637/c 530 bp mRNA EST 08-SEP-1999
 LOCUS 701676626 A. thaliana, Columbia Col-0, inflorescence-1 Arabidopsis
 DEFINITION thaliana cDNA clone 701676626, mRNA sequence.
 ACCESSION A1995637
 VERSION A1995637.1 GI:5842542
 KEYWORDS EST.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicot;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 530)
 Chan, D., Momiya, M., Chan, E., Mooney, M., Carroon, B., Gilliland, D.,
 Wang, X., Hillman, J., Guebler, K., Kim, C., Doyle, M., Brzoska, P.,
 Gorgone, G., Burns, D., Griffin, J., Mouanoutoua, M., Nguyen, D., Tan, R.,
 Rose, M., Warren, B., Ton, B., Kastury, K., Borillo, C., Carpio, T.,
 Policky, J., Suzuki, G., Argentine, C., Shah, S., Nobrega, A., Murry, L.,
 Turner, C., Krikorian, S., Elder, L. and Hanson, D.
 Arabidopsis thaliana Gene Expression MicroArray
 Unpublished (1999)
 JOURNAL Contact: David Smoller, Ph.D.
 COMMENT Genome Systems, Inc., a wholly owned subsidiary of Incyte
 Pharmaceuticals, Inc.
 4633 World Parkway Circle, St. Louis, MO 63134, USA
 Tel: 877-577-2733
 Fax: 314-427-3324
 Email: service@genomesystems.com.

FEATURES

source

1..530
 Location/Qualifiers
 /organism="Arabidopsis thaliana"
 /cultivar="Columbia Col-0"
 /db_xref="taxon:3702"
 /clone="701676626"
 /clone_1b="A. thaliana, Columbia Col-0, inflorescence-1"
 /tissue="inflorescence"
 /dev_stage="4 - 7 weeks"
 /note="Vector: pSPORT; Site_1: NotI; Site_2: SalI; cDNA library was derived from untreated inflorescence tissue from Arabidopsis thaliana, Columbia Col-0, at 4 - 7 weeks. Plants were grown in 1:1:1 peat moss/vermiculite/perlite soil at 22 deg. C +/- 3 deg. C under constant light, and watered with fertilizer. cDNA synthesis was initiated using a NotI-oligo(dT) primer. Double-stranded cDNA was blunt-ended, ligated to SalI adaptors, digested with NotI, size-selected, and cloned into the NotI and SalI sites of the pSPORT vector."
 BASE COUNT 163 a 108 c 100 g 158 t 1 others
 ORIGIN

Query Match 85.5%; Score 18.8; DB 10; Length 530;
 Best Local Similarity 90.9%; Pred. No. 1.1e+02;
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ttttgatcacaaagtctgata 22
 ||||| ||||| ||||| |||||
 Db 254 TGTGTGTTCCACAAGATTGACA 233

RESULT 6
 BE977618/c 608 bp mRNA EST 04-OCT-2000
 LOCUS bs66h08.y1 Drosophila melanogaster adult testis library Drosophila
 DEFINITION melanogaster cDNA clone bs66h08 5', mRNA sequence.
 ACCESSION BE977618
 VERSION BE977618.1 GI:10608272
 KEYWORDS EST.
 SOURCE fruit fly.
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 608)
 Andrews, J., Bouffard, G. and Oliver, B.
 Drosophila melanogaster testis expressed sequence tags
 Unpublished (1999)
 COMMENT Contact: Brian Oliver
 Laboratory of Cellular and Developmental Biology
 MIDDK, National Institutes of Health
 6 Center Drive MSC 2715, Bldg 6, Rm B1-13, Bethesda, MD 20892 USA
 Fax: (301) 496 5239
 Email: oliver@helix.nih.gov,
 http://www.niddk.nih.gov/intram/people/bolliver.htm
 Tissue isolation and library construction performed at the National
 Institute of Diabetes and Digestive and Kidney Diseases, NIH (see
 http://www.niddk.nih.gov/intram/people/bolliver.htm). DNA sequencing
 and analyses performed by National Institutes of Health Intramural
 Sequencing Center (NISC; see http://www.nisc.nih.gov).
 Plate: 66 row: h column: 08
 Seq primer: M13Rpl reverse primer (ABI).
 Location/Qualifiers

FEATURES

source

1..608
 Location/Qualifiers
 /organism="Drosophila melanogaster"
 /strain="y[*] w[67c1]/y"
 /db_xref="taxon:7227"
 /clone="bs66h08"
 /clone_1b="Drosophila melanogaster adult testis library"
 /sex="male"
 /dev_stage="1-5 day adult"
 /lab_host="SOLR (Stratagene)"
 /note="Organ: testis; Vector: pBluescript SK (Stratagene); Site_1: EcoR I; Site_2: Xho I; Testes dissected from 1-5 day adult y[*] w[67c1]/y males raised at 25oc. RNA isolated using Trizol (Life Technologies) and a single round of Poly(A)+ selection using Oligotex (Qiagen). cDNA library constructed using Stratagene Zap-cDNA synthesis kit. Oligo dt-primed, size fractionated ~1-6 kb, and directionally cloned at EcoRI and XhoI in Uni-Zap XR. Following a single round of amplification pBluescript SK phagemids were mass excised. A distribution channel for clones is being sought, but not currently available. Requests for clones cannot be honored."
 BASE COUNT 173 a 146 c 161 g 128 t
 ORIGIN

Query Match 83.6%; Score 18.4; DB 11; Length 608;
 Best Local Similarity 95.0%; Pred. No. 1.7e+02;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 ttggatcacaaagtctgata 22
 ||||| ||||| ||||| |||||
 Db 74 TTGGATCCCAAGATTGATA 55

RESULT 7
 CNS02BWA/c 934 bp DNA GSS 12-MAY-2000
 LOCUS Tetraodon nigroviridis genome survey sequence PUC-Orl end of clone
 DEFINITION 254C22 of library G from Tetraodon nigroviridis, genomic survey

FEATURES
source

Location/Qualifiers
1. .633
/organism="Medicago truncatula"
/cultivar="genotype A17"
/db_xref="taxon:3880"
/clone="pD5LC-40G21"
/clone_lib="DSLC"
/tissue_type="leaves and cotyledons"
/dev_stage="mixture of cotyledons from five days old plants and leaves obtained from two weeks old plants"
/lab_host="E. coli strain SOLR"
/note="Vector: pBluescript SK +/-; Site_1: EcoRI; Site_2: XhoI; cDNA was prepared from polyA+ enriched RNA from the mixture of cotyledons of five days old plants and leaves of two weeks old plants. The cDNA was directionally ligated into the Uni-ZAP XR vector from Stratagene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-ZAP phase using Ex-Assist helper phage and propagated in SOLR cells."

BASE COUNT 148 a 132 c 124 g 229 t

ORIGIN

Query Match 80.9%; Score 17.8; DB 11; Length 633;
Best Local Similarity 90.5%; Pred. No. 3.2e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 tcttgatcacacagattgat 21
|||||
Db 188 TGTTCATCACAAGATAGAT 168

RESULT 10
LOCUS BB288727 286 bp mRNA EST 09-JUL-2000
DEFINITION BB288727 RIKEN full-length enriched, 2 cells egg Mus musculus cDNA
clone B020032C12 3', mRNA sequence.
BB288727
BB288727.1 GI:8989176

KEYWORDS
SOURCE
ORGANISM

house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 286)

REFERENCE
AUTHORS

Kono, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T., Carninci
'P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N.,
Hirozane, T., Hori, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M.,
Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N.,
Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C., Kusakabe, M.,
Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y.,
Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata, Y.,
'Y., Shigemoto, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Sugihara, Y.,
'Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tomihaga, N., Toya, T.,
'Tsunoda, Y., Watahiki, A., Watanabe, S., Yamamura, T., Yamanaka, I.,
'Yano, R., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino, M.,
'Muramatsu, M. and Hayashizaki, Y.

TITLE
JOURNAL
COMMENT

Unpublished (2000)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoaka, S., Sasaki, N.,
'Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Thermolabile and thermolabile enzymes by
trehalose and its application for the synthesis of full length

FEATURES
source

CDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh, M., Katsunai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,
Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y.,
and Hayashizaki, Y.
Automated filtration-based high-throughput plasmid preparation
system. Genome Res. 9 (5), 463-470 (1999)
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
19-44 (1999)
Please visit our web site (<http://genome.rtc.riken.go.jp>) for
further details.

BASE COUNT
ORIGIN

117 a 66 c 20 g 83 t

Query Match 79.1%; Score 17.4; DB 10; Length 286;
Best Local Similarity 94.7%; Pred. No. 4.3e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 gtggatcacacagattga 20
|||||
Db 169 GTTGATCACAATATTGA 151

RESULT 11
LOCUS BE581720/c 407 bp mRNA EST 09-MAY-2001
DEFINITION BE581720 kg51c07.y1 TB995TM-SSR strongyloides stercoralis cDNA 5' similar to
WP:119810.2 CE16413 ;, mRNA sequence.
BE581720
BE581720.1 GI:9832662

KEYWORDS
SOURCE
ORGANISM

Strongyloides stercoralis.
Strongyloides stercoralis.
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida;
Panagrolaimidae; Strongyloidiidae; Strongyloides.

REFERENCE
AUTHORS

McCartter, J., Clifton, S., Chiapelli, B., Pape, D., Martin, J., Wylie, T.,
Dante, M., Maria, M., Hillier, L., Kucaba, T., Theising, B., Bowers, Y.,
'Gibbons, M., Ritter, E., Bennett, J., Franklin, C., Tsagaris, V., R.,
'Ronko, I., Kennedy, S., Maguire, L., Beck, C., Underwood, K., Steptoe, M.,
'Allen, M., Person, B., Swaller, T., Harvey, N., Schuck, R., Kohn, S.,
'Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and
Wilson, R.

The Washington Univ. Nematode EST Project, 1999
Unpublished (1999)
Contact: McCartter JP
The Washington Univ. Nematode EST Project, 1999
Washington University School of Medicine

Query Match 79.1%; Score 17.4; DB 10; Length 646;

Best Local Similarity 94.7%; Pred. NO. 4.9e+02;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 ttgatcacagaattgat 21
|||||

DB 256 TTGATCAGTACTGATTGAT 238

RESULT 14
BG448853 671 bp mRNA EST 16-MAR-2001

LOCUS NF003C10IN1F1082 Insect herbivory Medicago truncatula cDNA clone

DEFINITION NF003C10IN 5', mRNA sequence.

ACCESSION BG448853

VERSION BG448853.1 GI:13367634

KEYWORDS EST.

SOURCE barrel medic.

ORGANISM Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae; Medicago.

1 (bases 1 to 671)

REFERENCE 1 (bases 1 to 671)
Korth, K., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J., Flores

AUTHORS H.R., Iman, J.T., Meller, J.W. and May, G.D.

Expressed Sequence Tags from the Samuel Roberts Noble Foundation

Medicago truncatula insect herbivory library

UNPUBLISHED (2000)

CONTACT: Korth K

Dept. of Plant Pathology

University of Arkansas

217 Plant Science Building, Fayetteville, AR 72701, USA

Tel: 501 575 5191

Fax: 501 575 7601

Email: kthorth@comp.uark.edu

Insert length: 671 Std Error: 0.00

Plate: 003 Row: C Column: 10

Seq primer: TCACACAGGAACTATGAC.

Location/Qualifiers

1. 671

/organism="Medicago truncatula"

/db_xref="taxon:3880"

/clone="NF003C10IN"

/clone_lib="Insect herbivory"

/tissue_type="local and systemic leaves"

/dev_stage="mature"

/note="Vector: Lambda Zap; Library was produced from fully

expanded M. truncatula leaves of plants fed upon by

Spodoptera exigua (beet armyworm) for 24 hours. Systemic

(undamaged leaves from injured plants) and wounded leaves

were harvested and pooled."

BASE COUNT 186 a 151 c 131 g 203 t

ORIGIN

Query Match 79.1%; Score 17.4; DB 11; Length 671;

Best Local Similarity 94.7%; Pred. NO. 5e+02;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 ttgatcacagaattgat 21
|||||

DB 252 TTGATCAGTACTGATTGAT 234

RESULT 15

BE283043

LOCUS BE283043

DEFINITION 601101323F1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:3493829 5',

mrna sequence.

ACCESSION BE283043

VERSION BE283043.1 GI:9159023

KEYWORDS

SOURCE EST.

ORGANISM house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 762)

REFERENCE NIH-MGC <http://mgs.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

UNPUBLISHED (1999)

CONTACT: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:

<http://image.llnl.gov>

Plate: LLM8541 row: n column: 06

High quality sequence stop: 661.

Location/Qualifiers

1. 762

/organism="Mus musculus"

/strain="CZECH II (fetal)"

/db_xref="taxon:10090"

/clone="IMAGE:3493829"

/clone_lib="NCI_CGAP_Lu29"

/tissue_type="spontaneous tumor, metastatic to mammary.

stem cell origin"

/lab_host="DH10B"

/note="Organ: Lung; Vector: PCMV-SPORT6; Site:1; Salt:

Site:2: NotI; Cloned unidirectionally. primer: oligo dt.

Library constructed by Life Technologies. Investigator

providing samples: Gilbert Smith, NIH"

BASE COUNT 193 a 172 c 207 g 190 t

ORIGIN

Query Match 79.1%; Score 17.4; DB 10; Length 762;

Best Local Similarity 94.7%; Pred. NO. 5.1e+02;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 tggatcacagaattg 19
|||||

DB 692 TGTGGATCAGCAACTTG 710

Search completed: December 15, 2001, 02:33:49
Job time: 4989 sec

Mon Dec 17 07:48:52 2001

us-09-380-826a-7.rst

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 01:19:15 ; Search time 7962.77 Seconds
(without alignments)
3250.941 Million cell updates/secTitle: US-09-380-826A-1
1477
Sequence: 1 gatactgcacgaactaac.....ccgtaaatcgtctctcag 1477Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 17159718 seqs, 8763200856 residues

Total number of hits satisfying chosen parameters: 34319436

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_NA_Main:*

```
1: /cgn2_6/ptodata/2/pna/US086_COMB.seq:*
2: /cgn2_6/ptodata/2/pna/US06_COMB.seq:*
3: /cgn2_6/ptodata/2/pna/US07_COMB.seq:*
4: /cgn2_6/ptodata/2/pna/US08_COMB.seq:*
5: /cgn2_6/ptodata/2/pna/US081_COMB.seq:*
6: /cgn2_6/ptodata/2/pna/US082_COMB.seq:*
7: /cgn2_6/ptodata/2/pna/US083_COMB.seq:*
8: /cgn2_6/ptodata/2/pna/US084_COMB.seq:*
9: /cgn2_6/ptodata/2/pna/US085_COMB.seq:*
10: /cgn2_6/ptodata/2/pna/US086_COMB.seq:*
11: /cgn2_6/ptodata/2/pna/US087_COMB.seq:*
12: /cgn2_6/ptodata/2/pna/US088_COMB.seq:*
13: /cgn2_6/ptodata/2/pna/US089_COMB.seq:*
14: /cgn2_6/ptodata/2/pna/US090_COMB.seq:*
15: /cgn2_6/ptodata/2/pna/US091_COMB.seq:*
16: /cgn2_6/ptodata/2/pna/US092_COMB.seq:*
17: /cgn2_6/ptodata/2/pna/US093_COMB.seq:*
18: /cgn2_6/ptodata/2/pna/US094_COMB.seq:*
19: /cgn2_6/ptodata/2/pna/US095A_COMB.seq:*
20: /cgn2_6/ptodata/2/pna/US095B_COMB.seq:*
21: /cgn2_6/ptodata/2/pna/US095C_COMB.seq:*
22: /cgn2_6/ptodata/2/pna/US095D_COMB.seq:*
23: /cgn2_6/ptodata/2/pna/US096A_COMB.seq:*
24: /cgn2_6/ptodata/2/pna/US096B_COMB.seq:*
25: /cgn2_6/ptodata/2/pna/US096C_COMB.seq:*
26: /cgn2_6/ptodata/2/pna/US096D_COMB.seq:*
27: /cgn2_6/ptodata/2/pna/US096E_COMB.seq:*
28: /cgn2_6/ptodata/2/pna/US097A_COMB.seq:*
29: /cgn2_6/ptodata/2/pna/US097B_COMB.seq:*
30: /cgn2_6/ptodata/2/pna/US097C_COMB.seq:*
31: /cgn2_6/ptodata/2/pna/US098_COMB.seq:*
32: /cgn2_6/ptodata/2/pna/US099_COMB.seq:*
33: /cgn2_6/ptodata/2/pna/US6000_COMB.seq:*
34: /cgn2_6/ptodata/2/pna/US6001_COMB.seq:*
35: /cgn2_6/ptodata/2/pna/US6002_COMB.seq:*
36: /cgn2_6/ptodata/2/pna/US6003_COMB.seq:*
37: /cgn2_6/ptodata/2/pna/US6004_COMB.seq:*
38: /cgn2_6/ptodata/2/pna/US6005_COMB.seq:*
39: /cgn2_6/ptodata/2/pna/US6006_COMB.seq:*
40: /cgn2_6/ptodata/2/pna/US6007_COMB.seq:*
41: /cgn2_6/ptodata/2/pna/US6008_COMB.seq:*
42: /cgn2_6/ptodata/2/pna/US6009_COMB.seq:*
43: /cgn2_6/ptodata/2/pna/US6010_COMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1477	100.0	1477	17	US-09-380-826A-1
2	814.2	55.1	1536	13	US-08-642-229-1
3	814.2	55.1	1536	13	US-08-965-623-1
4	810.2	54.9	1508	17	US-09-198-935-14
5	810.2	54.9	1508	17	US-09-339-159-34
6	808.2	54.7	1508	27	US-09-694-531-14
7	808.2	54.7	1508	27	US-09-694-531-14
8	803.6	54.4	1501	29	US-09-745-476-1
9	803.2	54.4	1501	29	US-09-745-476-1
10	803.2	54.4	1501	29	US-09-745-476-1
11	803.2	54.4	1501	29	US-09-745-476-1
12	803.2	54.4	1501	30	US-09-791-592-1
13	803.2	54.4	1501	30	US-09-791-592-1
14	803.2	54.4	1501	31	US-09-821-016-5
15	801.6	54.3	1555	9	US-08-520-946-160
16	801.6	54.3	1555	13	US-08-923-030-11
17	801.6	54.3	1555	17	US-09-349-883-11
18	801.6	54.3	1555	54	US-06-219-361-21
19	801.6	54.3	1555	54	US-06-219-361-21
20	801.6	54.3	1555	54	US-06-219-361-21
21	801.6	54.3	1555	54	US-06-219-361-21
22	801.6	54.3	1555	54	US-06-219-361-21
23	801.6	54.3	1555	54	US-06-219-361-21
24	801.6	54.3	1555	54	US-06-219-361-21
25	801.6	54.3	1555	54	US-06-219-361-21
26	801.6	54.3	1555	54	US-06-219-361-21
27	801.6	54.3	1555	54	US-06-219-361-21
28	801.6	54.3	1555	54	US-06-219-361-21
29	801.6	54.3	1555	54	US-06-219-361-21
30	801.6	54.3	1555	54	US-06-219-361-21
31	801.6	54.3	1555	54	US-06-219-361-21
32	801.6	54.3	1555	54	US-06-219-361-21
33	801.6	54.3	1555	54	US-06-219-361-21
34	801.6	54.3	1555	54	US-06-219-361-21
35	801.6	54.3	1555	54	US-06-219-361-21
36	801.6	54.3	1555	54	US-06-219-361-21
37	801.6	54.3	1555	54	US-06-219-361-21
38	801.6	54.3	1555	54	US-06-219-361-21
39	801.6	54.3	1555	54	US-06-219-361-21
40	801.6	54.3	1555	54	US-06-219-361-21
41	801.6	54.3	1555	54	US-06-219-361-21

42 789.8 53.5 1516 18 US-09-426-633A-1 Sequence 1, Appl 1
43 789.8 53.5 1516 18 US-09-426-868-1 Sequence 1, Appl 1
44 785.4 53.2 21083 39 US-60-061-998-618 Sequence 618, App
45 783.6 53.1 1515 29 US-09-726-774-9 Sequence 9, Appl 1

ALIGNMENTS

RESULT 1
US-09-380-826a-1
Sequence 1, Application US/09380826a
GENERAL INFORMATION:
APPLICANT: Chappel, Rod
TITLE OF INVENTION: LEPTOSPIRA PATHOGENS
FILE REFERENCE: DATE: 79-001APC
CURRENT APPLICATION NUMBER: US/09/380-826A
CURRENT FILING DATE: 1999-11-22
PRIOR APPLICATION NUMBER: PCT/AU98/00145
PRIOR FILING DATE: 1998-03-06
PRIOR APPLICATION NUMBER: AU P05494/97
PRIOR FILING DATE: 1997-03-07
NUMBER OF SEQ ID NOS: 26
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 1477
TYPE: DNA
ORGANISM: Leptospira fainei
US-09-380-826a-1

Query Match 100.0%; Score 1477; DB 17; Length 1477;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1477; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gatcatgctcagaactaacgtctggcgcgctcttaaacatgcaagtcgagcggttag 60
DB 1 gatcatgctcagaactaacgtctggcgcgctcttaaacatgcaagtcgagcggttag 60
QY 61 caataccctgacgagcgagcggtgtagtaacacgtgttaacatgcaagtcgagcggttag 120
DB 61 caataccctgacgagcgagcggtgtagtaacacgtgttaacatgcaagtcgagcggttag 120
QY 121 ctctccgaagaagaagtaataccgtagtccctgttgatcaacaagatttgaatgtaa 180
DB 121 ctctccgaagaagaagtaataccgtagtccctgttgatcaacaagatttgaatgtaa 180
QY 181 agattttatgtcttgagatgagcccgcgccgattagctagcttgtagagtaatgctca 240
DB 181 agattttatgtcttgagatgagcccgcgccgattagctagcttgtagagtaatgctca 240
QY 241 ccaagcgacgagtcggtgacgagcgaggtgtcgcgcacaaatggaactgagacac 300
DB 241 ccaagcgacgagtcggtgacgagcgaggtgtcgcgcacaaatggaactgagacac 300
QY 301 ggtlccatctctactcgaaggagcgagcagtttaagaatctgtctcaatgvgggaaacctgaa 360
DB 301 ggtlccatctctactcgaaggagcgagcagtttaagaatctgtctcaatgvgggaaacctgaa 360
QY 361 gcaagcgacgagtcggtgaggaagaaggtcttgagattgtaaaattatagcagaagaaa 420
DB 361 gcaagcgacgagtcggtgaggaagaaggtcttgagattgtaaaattatagcagaagaaa 420
QY 421 ataagcagcaatgtgtagtgaactgcttaagaacgcgcgttaactagctgcagcagc 480
DB 421 ataagcagcaatgtgtagtgaactgcttaagaacgcgcgttaactagctgcagcagc 480
QY 481 cgcggttaatacgtatggtgcaagcgttcttgcgaatcatctggcgtaaggggtgctag 540
DB 481 cgcggttaatacgtatggtgcaagcgttcttgcgaatcatctggcgtaaggggtgctag 540
QY 541 cggattgtgaagtcaggtggaagaactcggtggtcaaccgttgctgcactggaacta 600
DB 541 cggattgtgaagtcaggtggaagaactcggtggtgcaaccgttgctgcactggaacta 600

DB 541 cggattgtgaagtcaggtggaagaactcggtggtcaaccgttgctgcactggaacta 600
QY 601 caagtcgagagtcgtggaagagcgcaagtggatcccaagtgtagcgtgaaatgctagat 660
DB 601 caagtcgagagtcgtggaagagcgcaagtggatcccaagtgtagcgtgaaatgctagat 660
QY 661 atctggaagacacgaatgagcgaagcgacttgcgtctaaactgacgtgagcag 720
DB 661 atctggaagacacgaatgagcgaagcgacttgcgtctaaactgacgtgagcag 720
QY 721 aaagcgtggttagtgaacccggtattagatcccggtatccacgcccctaaactgttctca 780
DB 721 aaagcgtggttagtgaacccggtattagatcccggtatccacgcccctaaactgttctca 780
QY 781 ccaagtcgtgggttttaaccctcaatgaacgaactgaagattagtagacgcctg 840
DB 781 ccaagtcgtgggttttaaccctcaatgaacgaactgaagattagtagacgcctg 840
QY 841 ggaatctgctcgaagaagtgaactcaaaaggaattgacgggggttcgcacaaagcgttga 900
DB 841 ggaatctgctcgaagaagtgaactcaaaaggaattgacgggggttcgcacaaagcgttga 900
QY 901 gcaatggttttaattcgaatgataccccaacaaactcaactcgtgggttgacatgtgaa 960
DB 901 gcaatggttttaattcgaatgataccccaacaaactcaactcgtgggttgacatgtgaa 960
QY 961 tcaatgagagataatgagccttcggcgagattcacaggtgtcatggtgtgtcgtcagc 1020
DB 961 tcaatgagagataatgagccttcggcgagattcacaggtgtgtcatggtgtgtcgtcagc 1020
QY 1021 tctgtgtcgtgagatggttgaagttcccgcaagcgacccctatgtagtgtct 1080
DB 1021 tctgtgtcgtgagatggttgaagttcccgcaagcgacccctatgtagtgtct 1080
QY 1081 acctaaagttggcaactggtacgaactcgcgtgacaaacgggaaggagcggtgaa 1140
DB 1081 acctaaagttggcaactggtacgaactcgcgtgacaaacgggaaggagcggtgaa 1140
QY 1141 cgtcaaatccatagtgcccttaattgtccagggccacacagtgctacaaatggtcgatcac 1200
DB 1141 cgtcaaatccatagtgcccttaattgtccagggccacacagtgctacaaatggtcgatcac 1200
QY 1201 aggttcgacaaactcgcaagaggaagtaatactctaaagtcgtgccagtcggattgg 1260
DB 1201 aggttcgacaaactcgcaagaggaagtaatactctaaagtcgtgccagtcggattgg 1260
QY 1261 gtcgtcaactcgaccccaatgaagtgcgaatcgctagtaatcgcgatcagatgcgcgg 1320
DB 1261 gtcgtcaactcgaccccaatgaagtgcgaatcgctagtaatcgcgatcagatgcgcgg 1320
QY 1321 tgaatacgttcccggaactgttacaacacgcgcgttcacacacacacgtgagttgggagcacc 1380
DB 1321 tgaatacgttcccggaactgttacaacacgcgcgttcacacacacacgtgagttgggagcacc 1380
QY 1381 gaagtcgtcttcttgaacgtaagagacagactaagtgtgaacccgttaaaagggtgt 1440
DB 1381 gaagtcgtcttcttgaacgtaagagacagactaagtgtgaacccgttaaaagggtgt 1440
QY 1441 gaagtcgttaacaaggtaacgttaacatcgatctcgtcag 1477
DB 1441 gaagtcgttaacaaggtaacgttaacatcgatctcgtcag 1477

RESULT 2
US-08-642-229-1
Sequence 1, Application US/08642229
GENERAL INFORMATION:
APPLICANT: Herwig, Russell P.
APPLICANT: Bielefeldt, Angela R.
APPLICANT: Stensel, H. David
APPLICANT: Strand, Stuart E.
TITLE OF INVENTION: Degradation of Environmental Toxins by a Filamentous Bacterium

```

: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Christensen O'Connor Johnson and Kindness PLLC
: STREET: 1420 Fifth Avenue, Suite 2800
: CITY: Seattle
: STATE: Washington
: COUNTRY: USA
: ZIP: WA 98101-2347
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/642,229
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/246,865
: FILING DATE: 20-MAY-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Sheiness, Diana K.
: REGISTRATION NUMBER: 35,356
: REFERENCE/DOCKET NUMBER: UOFW19233
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (206) 682 8100
: TELEFAX: (206) 224 0779
: TELEX: 4938023
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1536 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: nucleic acid
: DESCRIPTION: "16S ribosomal DNA"
: HYPOTHEetical: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: Commomonas testosteroni ATCC No. 11996
: US-08-642-229-1

Query Match          55.1%; Score 814.2; DB 10; Length 1536;
Best Local Similarity 74.7%; Pred. No. 3.5e-251;
Matches 1113; Conservative 0; Mismatches 343; Indels 33; Gaps 6;

```

```

Db 374 AAAGCTGATCCAGCAATGCCGCTGCAGATGAAGAGCCCTCGGGTTGTAACCTGCTTT 433
Qy 411 ggcaggaanaataaga-----gcaatgttgttgttactgtccta--aag 455
Db 434 GTACGGAAAGAAAGCCTGGGCTAATATCCCGGGCTGACGCTACCGTAAGATAAG 493
Qy 456 caccgctactacgtgtccagccggttaatacgtatgtgtcaagctgttcgtga 515
Db 494 CACCGCTAACTACGTGCGCAGACGCGCGGTAACTAGGTGGGTGCAAGCTTAATCGGA 553
Qy 516 tcaatggcgtaaaagggtgcgttagcggaattgttaagtcagtgltgaaaactgcggctc 575
Db 554 TTACTGGGCTAAAGCGCTGCCAGCGCGGTGTTGTAGACAGTGTGAATCCCGGGCTC 613
Qy 576 aaccgctgcctgcacttgaactaagaatctcggaattgggagagcaagtggaattcc 635
Db 614 AACCTGGGAAGTCCCTTTGTGACTGCAGAGCTAGAATGGCGGAGGGGATGAATTC 673
Qy 636 aggtgtagcggtgaatgtcgtatgatalctcgaaggaaaccagtggtgcgaagcactgtct 695
Db 674 GCGTGTACGACGTAATGCGGTAGATGTGGGAGGAACACCGATGGGAGGCAATCCCT 733
Qy 696 ggtcaaaactgaacgtctgaagcagaagcgttggtagtaaaacggaatagatacccg 755
Db 734 GGGCTGACATGACGCTCATGCGAAGCGTGGGAGGAGCAACAGAGATTAGATACCTCG 793
Qy 756 taatcagccctaaacgttgttctaccagttgttgggggtttaaaccctcagtaagaac 815
Db 794 TATGTCCACGCCCTTAAACGATGTCACACTGGTGTGGGTCTTAACCTGACCTGATTAAG 853
Qy 816 ctaagcgttaagtagaaccccgctgggactatgctcgaagagtgaactcaagaat 875
Db 854 CTATCCGCTGAAGTTATACCGCTGGGAGTACGCGCGCAAGGTGAACCTCAAGGAATT 913
Qy 876 gacgggggtccgcacaaagcgttgaagcatgtgttaacttcgatgataccaccaaacct 935
Db 914 GACGGGGAGCCCGACAAAGCGGTGATGATGGTTAATTCGATCGAACGCGAANAACCT 973
Qy 936 cacttggttgacatgatacgtatcgaatcatgtagagatatatgaccttgggcaga---- 991
Db 974 TACCCACCTTTGACATGGAGAGAACTTACCAAGATGTTGTGCTGCCAAGAGAACT 1033
Qy 992 -ttcaaggtgtcgtcagttgttcgtcaagctcgttgcgtgaagatgttgggttaagtccg 1050
Db 1034 GCACACAGGTGCGTACATGCGTGTGCTGCTGACGTGTCGTAAGATGTTGGGTTAAGTCCCG 1093
Qy 1051 caagagcgcacaaccctatcgtatgttgccttaagtgttgacactgttgaacaaactgc 1110
Db 1094 CAAACGAGCCCAACCTTGGCATTTAGTTGCTACATTCAGTTGAGCACCTCTAATGGGACTGC 1153
Qy 1111 cgttgaacaaacggaggaagcggggaatgacgtcaaatcctcatgtgaccttatgtccag 1170
Db 1154 CGGTGACAAACCGAGAGAGGTGGGATACGTAAGTCTCTATGGCCCTTATAGTGGG 1213
Qy 1171 gccacacacgtgtcacaatggtccgatagagagtggtcgaactcgcgaagaagagagtaat 1230
Db 1214 GCTACACAGCTCATTAATAATGCGTGTGTAACAAAGGTTGCCAACCCGGAGGGAGCTAAT 1273
Qy 1231 ctctaaagtcggtccagttcgaattgggtgtcgaactcgaccctcgaatgaatcgaat 1290
Db 1274 CCCATTAAACCGAATCGTACGATCCGATCGCATTCGACCTCGATCGAAGTCCGAAT 1333
Qy 1291 cgttagtaatcgtgagtagcgaacgcgcgttgtaatacgttccggacctgtgaacacgc 1350
Db 1334 CGCTAATTAATCGGTGATCGAATGTACGCTGATACGTTCCGGGTCTTGTGTAACACCG 1393
Qy 1351 cccgtcacacacacgtatgggtgagcaccgaagtggtcttltgttaacgttaagagaca 1410
Db 1394 CCGGTACACACATGAGGAGGGGCTCGCCAGAAAGTATG-ATAGCTTAACGTTAAGAGGGGC 1452
Qy 1411 gactactaaggtgaaaactcgtlaaagggttgaagtcgtatacaagtaacc 1459
Db 1453 GCTTACCACGGGCGGGTCTGCTGACTGGGGTGAAGTGTGTAACAAAGGTATGC 1501

```

```

RESULT 3
US-08-965-623-1
: Sequence 1, Application US/08965623
: GENERAL INFORMATION:
: APPLICANT: Bielefeldt, Angela R.
: APPLICANT: Stensel, H. David
: APPLICANT: Strand, Stuart E.
: APPLICANT: Herwig, Russell P.
: TITLE OF INVENTION: Degradation of Environmental Toxins by a
: TITLE OF INVENTION: Filamentous Bacterium
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Christensen O'Connor Johnson and Kindness PLLC
: STREET: 1420 Fifth Avenue, Suite 2800
: CITY: Seattle
: STATE: Washington
: COUNTRY: USA
: ZIP: WA 98101-2347
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION NUMBER: US/08/965,623
: APPLICATION NUMBER: US 08/642,229
: FILING DATE: November 6, 1997
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/246,865
: FILING DATE: 20-MAY-1994
: APPLICATION NUMBER: US 08/599,867
: FILING DATE: 12-FEBRUARY-1996
: APPLICATION NUMBER: US 08/642,229
: FILING DATE: 30-APRIL-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Sheiness, Diana K.
: REGISTRATION NUMBER: 35,356
: REFERENCE/DOCKET NUMBER: UOF111488
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (206) 682 8100
: TELEFAX: (206) 224 0779
: TELEX: 4938023
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1536 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: nucleic acid
: DESCRIPTION: "16S ribosomal DNA"
: HYPOTHEICAL: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: Comamonas testosteroni AFCC No. 11996
: US-08-965-623-1

```

```

Query Match 55.1%; Score 814.2; DB 13; Length 1536;
Best Local Similarity 74.7%; Pred. No. 3, Se-251;
Matches 1113; Conservative 0; Mismatches 343; Indels 33; Gaps 6;

```

```

Db 136 TGGGGATTAACCTACGAAAGTAGTAATACCGCATGAGATCTACGATGAAGA - 193
Qy 171 tgatagtgaaagattatgttgtagatagcccgccgtagttagtgtagg 230
Db 194 GGGGACCTTCGGGCTTGCTGCTACTAGAGCGGCTGATGCGAGATTAGATGTGTGGG 253
Qy 231 taatgctcaccagagcgacatcgtagccgcttgaaagagtgctcgacacaatga 290
Db 254 TAAAGGCTTACCAAGCTCGCATCTGTAGCTGTGTGAGAGACGACGACACACTGG 313
Qy 291 actgaagacaggtccatactctcagagagagcagcagtlgaatctgtcgaatggg 350
Db 314 ACTGAGACACCGCCAGACTCTTACGGGAGCAGACAGTGCGGAATTGTGACAAATGGGCG 373
Qy 351 aaacctgaagcagcagcgccgctgtagcagaagaagtcctgattgtaagtctta 410
Db 374 AAGCTGATCCAGCAATGCCGCTGACAGATGAAGGCCCTCGGGTTGTAACTGCTTT 433
Qy 411 ggcagaaaaaataagca-----gcaatgtatgtactactccta--aag 455
Db 434 GTACGGAACGAAAGCCGCGGCTAATATCCCGGTCATGACGTAACGTAAGAAATAG 493
Qy 456 caccgctcactacgtgccaagcagcggtgaatagatgtagcaggtgtgtcgaa 515
Db 494 CACCGGCTAAGTACGAGTGCAGACGCGGTAATAGTAGGCTTAATCGAA 553
Qy 516 tcaatggcgtaaaaggtgctgtagcgatctgaagtcaggtgtgaaacatcgggctc 575
Db 554 TTACTGGCGTAAAGCGTCCGACGGGTTTGTAAAGCACTGGTAATCCCGGGCTC 613
Qy 576 aaccgtggtcgtcaactgaaactacaaagtcgtgaagtttggaagagcagttgaatcc 635
Db 614 AACCTGGCAACTGCATGTGCTGCTCAAGGCTAGAGTCCGCAAGGGGATGCAATTC 673
Qy 636 aggtgtgacggtgaaatgctgtagatcttgaggaaccccgatggcgaagcgactgtc 695
Db 674 GCGTGTACAGTGAATGCTGATATGCGGAGAACCCCATGGCGGAGCAATCCCT 733
Qy 696 ggtcctaaactgacgtctgagcagcaaaagcgtggtagtaaaagcgtatataaccgccg 755
Db 734 GGGCTCGACTGACGCTGCTATACAGAAAGCTGGGGACCAACAGGATTATATACCTCG 793
Qy 756 taatccacgcccctaaagctgtgtctaccagttgttggtgtttaaaccctcagtaaac 815
Db 794 TAGTCCACGCCCTTAACCATGTCAACTGTGTGGCTTAACGTCACTAGTAACGAG 853
Qy 816 ctacagatgaatgaagccgtgggactatgtctgcaagaagaaactaaaggaat 875
Db 854 CTAACGCGCTGAAGTTGACCGCTGGGAGTAGCGCCGCAAGGTTGAACCTAAAGGAAT 913
Qy 876 gacggggtccgcacaaagcgtgtgagcaatgtgttlaatcagatgaatacccaaaact 935
Db 914 GACGGGGACCCGCAAAAGCGGTGATGTGTTAATTCGATGCAACCGGAAACCACT 973
Qy 936 caacttggtctgacatgtagatctgatactatgtagagatatgagccttcggcaga---- 991
Db 974 TACCCACCTTGAACATGACAGAACTTAACAGAGATGGTTGGTCTCGAAAGAAACCT 1033
Qy 992 -ttcagaagtgatgcatggtgtgtcagcgtgtcgtgtgagatgttgtagtccg 1050
Db 1034 GCACACAGGTCTGCTGATGCTGTGCTGACGTGTCTGTGAGATGTGTGAGTAACTCCG 1093
Qy 1051 caacgagcgcaaccctatcgtatgtgtcactaaagtgtggcactgtgtagaanaatg 1110
Db 1094 CAACGAGCGCAACCTTGCTCATTTAGTTGCTCATTTAGCACTGTATGGAGACTGC 1153
Qy 1111 cgtgtgacaaacgggaagagcggtgtagatgaatccctcagtccttaatgtccag 1170
Db 1154 CGGTGTACAAACCGGAGGAGGTGGGATGAGAGCTCAAGTCCATGACGCTTATAGTGG 1213
Qy 1171 gccacacagctgtacaaatgtgcatacagaaggttcgcaactcgcgaagagaggaat 1230
Db 1214 GCTACACAGCTCATATACATGTGCTGTACAAAGGTTGCCAACCCGCGGAGGAGCTAAT 1273

```



```
QY 1231 ctctaaagtcggtcccaagtcggtatgtggtctgtcaactcgaaccatgaatcgaat 1290
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1274 CCATTAAGCACTCGTAGTCCGATCGCATCTCGACACTCGACTGCGTGAAGCGGAAT 1333
QY 1291 cgtagtaatcgtcgatcgaatcgtcggtgtaatacgttcccgacctgtgataccg 1350
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1334 CGCTAGTAATCGTGATCAAAATGTACGCTGAATAGTTCCCGGCTCTGTACACACCG 1393
QY 1351 ccgtacaccactgagtggtggaagcaccggaagtgtcttgttaacgttaaggagaa 1410
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1394 CCGCTACACACAGGAGCGCGGTCTCGCAGAAAGTAG-GTAGCCTTAACCGTAAGAGAGGC 1452
QY 1411 gacttaagtgtaaacctgttaaagggtgtaagtcgtaacgaagtacc 1459
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1453 GCCTACACGCGCGGTCTGTGACTGCGGTGAAGTCTGAACAGGTAGC 1501
```

```
RESULT 4
US-09-198-955-14
: Sequence 14, Application US/09198955
: GENERAL INFORMATION:
: APPLICANT: Andersen, Lene Nonboe
: APPLICANT: Schulein, Martin
: APPLICANT: Lange, Niels Erik Krebs
: APPLICANT: Bjoernvad, Mads Eskelund
: APPLICANT: Moeller, Soeren
: APPLICANT: Glad, Sanne O. Schroeder
: TITLE OF INVENTION: Novel Pectate Lyases
: FILE REFERENCE: 5378,200-US
: CURRENT APPLICATION NUMBER: US/09/198,955
: EARLIER FILING DATE: 1998-11-24
: EARLIER APPLICATION NUMBER: 1343/97
: EARLIER FILING DATE: 1997-11-24
: EARLIER APPLICATION NUMBER: 1344/97
: EARLIER FILING DATE: 1997-11-24
: EARLIER APPLICATION NUMBER: 60/067,249
: EARLIER FILING DATE: 1997-12-02
: EARLIER APPLICATION NUMBER: 60/067,240
: EARLIER FILING DATE: 1997-12-02
: EARLIER APPLICATION NUMBER: 09/073,684
: EARLIER FILING DATE: 1998-05-06
: EARLIER APPLICATION NUMBER: 09/184,217
: EARLIER FILING DATE: 1998-11-02
: NUMBER OF SEQ ID NOS: 14
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 14
: LENGTH: 1508
: TYPE: RNA
: ORGANISM: Bacillus sp.
US-09-198-955-14
```

```
Query Match 54.9% Score 810.2; DB 15; Length 1508;
Best Local Similarity 61.7%; Pred. No. 6,9e-250;
Matches 921; Conservative 220; Mismatches 303; Indels 49; Gaps 10;
```

```
QY 15 actaagcgtggcgcggtctttaacatgcaagtcgagcg-----gtag 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2 acgaacgcgugcgugcgucauaacgaucgaugcgagcgagacaauuagagcuugccc 61
QY 61 caatactacgagcgaacgggtgagtaacacgtgtgtaactcttcctcg-egtcggata 119
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 62 uaaauugagcgcgcguguguaacacgugugcaaccugcgaccugcuugagcuuguaa 121
QY 120 acttcggaagaagaactaataccgagtagtctctgtgtgataccaagattgtagta 179
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 122 acacagcgaagaacugugcuaaaccgtauaucug-aggaaucauaauncucuuuuaa 180
QY 180 aagat-----ttatgctgtgagatgagcccgcgccgattagctagtgtgaggt 231
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 181 aagagugcuccgcuacacuaacggaugugcccgcgcgcauuauguuuuguaaggu 240
```

```
QY 232 aatgctcacaaagcgacgacgtcgttagccggtcgtgagaggtgtccgcacaatgaa 291
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 241 aacgcuuaacaaagcgacgacgugcugacgcacgugagggugaugcgcacacuuagg 300
QY 292 ctgagacaggtccatctcctcaacggagcagcagtcgaatctgtcccatgggga 351
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 301 cugagacacgcccacgacucucacgagggagcagcaguaaggaucucuccgaugagca 360
QY 352 aacctgaagcagcgcgccgtgtaacgagaaggtcgtgtagtctgtaagtcattag 411
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 361 aagucugacggagcaaacgcgcgugaugaugaggguuuucgcuguaaagcucugug 420
QY 412 gcaaggaaaaaagcagcaatgtatgtatgtactgtcct-----aaag 455
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 421 uuaugagaacaaagcugcugucuaaauaggcgccacucugacgucuaacacgaag 480
QY 456 caaccgtaactacgtgcacgacgacggttaactgtagtgcgaagctgttcgaa 515
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 481 ccaacgcuacacugcugcagcagcgcuuaaauagugugcagcgugucgga 540
QY 516 tcatggcgctaaagggtgcgtgaggttgttaagtcagtgctgtaaaactgcggctc 575
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 541 uuaugugcguaaagcgcgacgagcgucuuuaagucugauguaaauccggguc 600
QY 576 aaccggtgcctgcacttgaaactacaaagtcgtgagtttggagaagcgaatgtcc 635
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 601 aaccgcgacgcuuauguaaauugaaucuguaagcgaagaagaguguaauucc 660
QY 636 agtgtgcgtgtgaatgtcgtatgtatctggaagaacacacagtcgcaaggcgtgtc 695
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 661 acgugugcguguaagcguagcuaugugagaaacacagugcgcaagcgcgucucu 720
QY 696 ggcctcaactgacgtgagcgaacgaagcggtgtgtagtaaacgggtatgatccccg 755
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 721 gducuguaacugacgucugagcgcgaaagcgugggagcgaacggauuugauaccug 780
QY 756 taatccacgccttaacacgtgtctaccagttgttgggggtttaa--ccctcagtaa 813
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 781 uagucacgcgcguuaacgaaugagugcuauguguuagggguuucgacccuagugcgca 840
QY 814 aacctacgattatgtacgcgcgtcggtggtactatgtcgcgaagaagttaactcaagaa 873
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 841 aguuuaacacuaaagcacucgcgcugggagagacgcagcgaaaguuuaacucaaagaa 900
QY 874 ttgacggggcttcgcacaaacggtgagacgtgtgttaatttcgtatcatcccaaaaac 933
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 901 uugacggggcgccgacaaagcagugagcauguguuuaucgaacacgcgcaagaac 960
QY 934 ctcaaccgtgtgacatgtaactgaaatcatgtagagata--tatgagccttcggcaga 991
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 961 cuuacagugucuaacuuacuuagaccuccuagagagaguuuuccuuggggaga 1020
QY 992 t---tcaacaggtctgtacgtgtgtcgcgtcgtgtcgtgtagatgttggtaacttc 1048
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1021 uaaugacagugugugcauguguuucgucagcugugcugagauuguguuuaauucc 1080
QY 1049 cgcacagcgcaaacccct-ctgctatgttccttaactgagtcgacgtgtgcaaac 1107
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1081 cgcacagcgcaaacccuugacuuuagugccagcauuaugugugcacaucuaagugac 1140
QY 1108 tgcggtgtaacaacgcgaggaagcgaggatgacgtcaaatctcatgtgccttaatc 1167
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1141 ugccguguaaacaacgcgaggaagugugagucuaaaucauacauaccuccuuaugac 1200
QY 1168 agggcacacacgctgtacaaatggtcgatcaagaggttcgcacaactgcgaagagagct 1227
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1201 ugggcuacacacgugcuaacauugauuacaaagacgaagaaacgcgcgugucgagc 1260
QY 1228 aatctcaaaagctgcctccagtcgtagtgggttcgcaactgcacccatgaaatgc 1287
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1261 aaucucaaaagccauucucuguuugaguuagcugcaacucugccuacuaagagcg 1320
QY 1288 aatcgtagtaatcgcggtacgacatgcgcggtgtaatcgttcccggaactgtataca 1347
```


Oy 1407 gacagactactaaagtgtgaaactcgtlaaagggtgtgaatcgtaacaagttacc 1459
| | | | : | | | : | | | | : | | | | | : |
Db 1441 gcuagccgcuaagugggacagaugaauuggugugaucguuaacaagaaguac 1493

RESULT 6

```

US-09-694-531-14
? Sequence 14, Application US/09694531
? GENERAL INFORMATION:
? APPLICANT: Andersen, Lene N.
? APPLICANT: Schultein, Martin
? APPLICANT: Lange, Niels E.
? APPLICANT: Bjornvad, Mads E.
? APPLICANT: Moller, Soren
? APPLICANT: Glad, Sanne O. S.
? APPLICANT: Kaupluehn, Markus S.
? APPLICANT: Schnorr, Kirk
? APPLICANT: Kongsbak, Lars
? TITLE OF INVENTION: Novel Pectate Lyases
? FILE REFERENCE: 5378.200-US
? CURRENT APPLICATION NUMBER: US/09/694,531
? PRIORITY FILING DATE: 2000-10-123
? PRIOR APPLICATION NUMBER: 09/198,955
? PRIOR FILING DATE: 1998-11-24
? PRIOR APPLICATION NUMBER: 1343/97
? PRIOR FILING DATE: 1997-11-24
? PRIOR APPLICATION NUMBER: 1344/97
? PRIOR FILING DATE: 1997-11-24
? PRIOR APPLICATION NUMBER: 60/067,249
? PRIOR FILING DATE: 1997-12-02
? PRIOR APPLICATION NUMBER: 60/067,240
? PRIOR FILING DATE: 1997-12-02
? PRIOR APPLICATION NUMBER: 09/073,684
? PRIOR FILING DATE: 1998-05-06
? PRIOR APPLICATION NUMBER: 09/184,217
? PRIOR FILING DATE: 1998-11-02
? NUMBER OF SEQ ID NOS: 32
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 14
? LENGTH: 1508
? TYPE: RNA
? ORGANISM: Bacillus sp.
US-09-694-531-14

```

Query Match	54.9%	Score 810.2;	DB 27;	Length 1508;
Best Local Similarity	61.7%	Pred. No. 6.9e-250;		
Matches 921;	Conservative 220;	Mismatches 303;	Indels 49;	Gaps 10

Oy 15 acaacgttcgcgcgctcttcaacatgcgaatcgcgcg-----gtag 60
 Db 2 acgaacgcgucgcgccgugccaaabaucgaagucgcgagacauuagagcugcucc 61
 Oy 61 caatactagtcgcgacgaacgggtgagtaacaatgtaactctcccg-agtctggagata 119
 Db 62 uaaauuguaatgcgcgacgcgugagaaacacguguggcgaacucgccugabagacuuaggaa 121
 Oy 120 acttcgcgaagaaggaaacttaactacgcgatagtcctgtctgataccaagatcttgatagta 179
 Db 122 acauacgagaauucgugucuaaueccgcgaatuaucug-aggaauugccaauaaccucuuagaa 180
 Oy 180 aagat-----ttatgtcttgagatgagccgcgacgcgatagcttggttgagtg 231
 Db 181 aagaugugucccgucuaacuaacuaacgugaugggccgcgagcauuatgacuauguguaaguu 240
 Oy 232 aatgcgtaccacaaagcgacgaatcgtagtcagccgcttgagaggtgtccgcgcacaaatgaa 291
 Db 241 aacggcuuaccaaagcgacgaugugugccgaccuugaggguguauggcgcacaaucungga 300
 Oy 302 ctgagacacggtccaatactctacggagagcagcttaagaatctgtctcaatggaggga 351
 Db 301 cugagatccgcgcacagacucuaacacggagagtgatgacgaagaaatuaacucgcgaalugagcga 360

[illegible]

Db 1441 gcuaagccgcuaagugugagacagaugauuggguguaugucuaacaagugac 1493

```
RESULT 7
US-60-279-526-1256
; Sequence 1256, Application US/60279526
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Id Groth
; APPLICANT: Sorokine, Alexei
; APPLICANT: Bolodine, Alexandre
; APPLICANT: Lapidus, Alla
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: 5963.000-US
; CURRENT APPLICATION NUMBER: US/60/279,526
; CURRENT FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 2232
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1256
; LENGTH: 5643
; TYPE: DNA
; ORGANISM: B. clausii
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(5643)
; OTHER INFORMATION: n = A,T,C or G
US-60-279-526-1256
```

Query Match 54.7%; Score 808.2; DB 60; Length 5643;
Best Local Similarity 75.2%; Pred. No. 5,5e-249;
Matches 1130; Conservative 0; Mismatches 328; Indels 45; Gaps 8;

Qy 1 gatcatgctcagaactaacgtctgacgctgcttaataatgcagtcgagcgg----- 55
Db 319 gatccctgctcagacgacgtgctgctgcttaataatgcagtcgagtcgacaga 378
Qy 56 -----ggtacaatacctcagcgcgaacgggtgagtaacagct-gtataatcct 105
Db 379 agggagctctgctccgcgacgttagcgcgacgggtgagtaacagctggtgcaactgcgc 438
Qy 106 ccgaatctcgagataactctccgaaagaaagctaaataccggaatgctcgtt-----g 158
Db 439 cttaagctcgagataactccggaacgcgagctaaataccggaatcctcctcaccct 498
Qy 159 gatcccaagattgtatgtaagattatattgctgtgagatgagccgcgcgacattgc 218
Db 499 ggagagaggggtgaagatggtctcgctacatcagggagatggtgcccgcgcgacattgc 558
Qy 219 tagttgtgagtaagtaatggtctcaccaagcgacatcggtagccgcgtgagagaggtgtcc 278
Db 559 tagttgtgagtaagtaatggtctcaccaagcgacatcggtagccgcgtgagagaggtgtcc 618
Qy 279 ggcacaatcgagactgagacacgctccatactcctcagggagagcgacgttaagaatctt 338
Db 619 ggcacactggtgagactgagacacgctccatactcctcagggagagcgacgttaagaatctt 678
Qy 339 gctcaatggtggaacccctggaagcgacgcgtgtaagaagaaagctctcgatgt 398
Db 679 ccgcaatggaagacgtctgacgagacgcgcgtggtgtaagaagacccctcgggtcgt 738
Qy 399 taagatcatgagcagaaaataaatacagcaatgtgatgtatgtaactgtcct----- 451
Db 739 taagatctgtgtgagggaaagaagcgtaacgttcgaatagggcggtacactgtcggtga 798
Qy 452 -----aaagcaacggtactaactagctgtccagcagccggtgtaactagtggtgcaa 502
Db 799 cctacacgaagcagcgctaaactagctgtccagcagccggtgtaactagtggtgcaa 858
Qy 503 ggcgtgtcgcgaatcattatgggtgtaaggtgtcgtaggggtttgttaagtatggtgta 562
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 859 ggcgtgtcgcgaatcattatgggtgtaaggtcgcgacgctccttaagtctgattgta 918
Qy 563 aaactgcggtctaacccgtgtgctgctgctgaaactaaactgattctggaagg 622
Db 919 aatccggggtcccaaccccgagcgccatctggaactctggagctctgagtcgaagaagg 978
Qy 623 caagtgtaatccaggtgtgtaacggtgaaatgcgtatagatatctggaggaacccagtcgcg 682
Db 979 agagtgtgattccacgctgtgtaacggtgaaatgcgtatagatatctggaggaacccagtcgcg 1038
Qy 683 aagcgactgtctgctgctaaactgacgctgagcgacgaaagcgctgggtatgaacggga 742
Db 1039 aagcgactgtctgctgctaaactgacgctgagcgacgaaagcgctgggtatgaacggga 1098
Qy 743 ttagatacccggttaactcacgcccctaaacgtgtctacacggtgtgtgggttttaa-- 800
Db 1099 ttagatacccggtgtgctgctgacgcccgttaaacggtgtgtgtgtgtgtgtgtgtgtgtgt 1158
Qy 801 cctctagtaacggaactaaacggtatgaacgctggtgggactatgtctgcgaagagtg 860
Db 1159 ccggtagtgccgaagttaacacatgaacactcgcctctgggagtgacgcccgaagctgt 1218
Qy 861 aaactcaaggaattgagcggtgtccgcacaagcggtgtgaaatggttttaattcagtg 920
Db 1219 aaactcaaggaattgagcggtgtccgcacaagcggtgtgaaatggttttaattcagtg 1278
Qy 921 ataccaccaaaactcactcgtgctgacatggaat---ctgaatcatgaagatatatg 977
Db 1279 caacgcgaagaactcactcgtgctgacatggaat---ctgaatcatgaagatatatg 1338
Qy 978 agccttcgggacagatcaggtgtgctgacatggtgtgtcagctcgtgtgtgagatgt 1037
Db 1339 ccttcggggtgcaaaagtgacaggtgtgtgcatgtgtgtgtcagctcgtgtgtgagatgt 1398
Qy 1038 ggtttatgtccgcgaacgagcgcaacccct-atcgtatgtgtcactccttaagtttgagac 1096
Db 1399 ggtttatgtccgcgaacgagcgcaacccct-atcgtatgtgtcactccttaagtttgagac 1458
Qy 1097 tgtgtaaacatctgcggtgtaacaaacggaagagcggtggaatgtaactaaatcctcatgtg 1156
Db 1459 tctaaagtgtacgtccggtgtaacaaacggaagagcggtggaatgtaactaaatcctcatgtg 1518
Qy 1157 ccttatgttccagggacacacacgctgtctacaatgtgcgatatagaaggtgtccaaactgc 1216
Db 1519 ccttatgttccagggacacacacgctgtctacaatgtgtgtacaatgtgtgcgaaacacgc 1578
Qy 1217 aagagggaggaatcttcaaaagtctgctccaggttcggtatgtgggtctgcaactcgacc 1276
Db 1579 gagggtgaagccaatcccataaagccatctcagttcggtatgtgcaggtcgcaactcgctgt 1638
Qy 1277 catgaagtcggaatcgtatgaatcgcggtatcagcatgtccggtgtgaatacgttccgga 1336
Db 1639 catgaagtcggaatcgtatgaatcgcggtatcagcatgtccggtgtgaatacgttccggt 1698
Qy 1337 cctgttacaacgcgccttccacacacactgagtggtgggaccccggaatgtgtctgttga 1396
Db 1699 tctgttacaacgcgccttccacacacactgagtggtgtgtgtgtgtgtgtgtgtgtgtgtgt 1757
Qy 1397 accgttaaggaagacactaactaaggtgaacactgtaaaaggggttgaatcgttaaaaggt 1456
Db 1758 accgttggagcagcgcgtcactaaggttggtgacaaatgattgggtgtaagtcgttaaaaggt 1817
Qy 1457 acc 1459
Db 1818 agc 1820

RESULT 8
US-09-680-598A-1291/c
; Sequence 1291, Application US/09680598A
; GENERAL INFORMATION:
; APPLICANT: Randy M. Berka
; APPLICANT: Id Groth Clausen

```

; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE OF INVENTION: Expression
; FILE REFERENCE: 10085.000-US
; CURRENT APPLICATION NUMBER: US/09/880,598A
; CURRENT FILING DATE: 2000-10-06
; NUMBER OF SEQ ID NOS: 1862
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 1291
; LENGTH: 2484
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: (1)...(2484)
; OTHER INFORMATION: n = A,T,C or G
US-09-680-598A-1291

Query Match          54.4%; Score 803.6; DB 27; Length 2484;
Best Local Similarity 75.4%; Pred. No. 1.2e-247;
Matches 1134; Conservative 0; Mismatches 324; Indels 46; Gaps 9;

QY 1 gatcatgctcagaactacgctgcgcgcgtcttaacatgcagtcgagcgggtag 60
DB 2180 GATCCTGCTCAGACGACGAGCTGGCGCTCTAATACATGCAAGTCGACGACCA 2121
QY 61 caatccct-----agcggcgaagcgttgatgaacagt-gtaatctctct 105
DB 2120 CGGGAGCTTTCCTCCTTAGCTAGCGGCGGAGCGGAGTACACGCTGGTAACCTGCCT 2061
QY 106 ccgaagctcgtgaactcctccgaagaagcctaataccgaagc-----ccgtgt 157
DB 2060 GTAAGACTGGGATACCTCCGGGAAACCGGGGCTAATACCGGATGCTGATTAACCGCAT 2001
QY 158 ggaatcaagaattgatagtaaatatattatgtcttgagatgagccgcggcagattag 217
DB 2000 GGTTCATCATTAAGAGTGGCTTTTAGTACCATCCATTACAGATGAGCCCGCGCATTAG 1941
QY 218 ctactgtgtgaagtaatgctaccagaagcgaagtcggtacccgcgtgaagaggtgtc 277
DB 1940 CTAGTTGGTAGAGTAAAGGCTACCAAGGCGAGCATGCTTACCGCATGGAAGAGGTGAT 1881
QY 278 cggcacaatggaacttgagacatgggccaaactcctcctacggaagcagcaagtaagaatct 337
DB 1880 CGGCACTGGGACTGAGAGAGCGGCCGACACTCTCTACGAGGAGCAGCAGTACGGGATGT 1821
QY 338 tgcctaatgagggaacccctgaagcagcgcgcgtgtgaacgaagaaggtctcgatt 397
DB 1820 TCCGCAATGAGAAAGTCTGACGAGCAACGCCGCTGATGATGAAGTTTTCGGATC 1761
QY 398 gtaagttcatatgagcagaagaataagaagaatgtagtgaactgtct----- 451
DB 1760 GTAATACTCTGTTTGTAGGAAAGAACAACTACCGTTGCAATAGGGGGGTAACCTTGACGCT 1701
QY 452 -----aaagacccgcctaactacgtgcgaagcgcgcgtgaataatcgtatgtca 501
DB 1700 ACCTAACCGAAGAACGACGCGCTAATACGTCGCCAGAGCCGGGTAATACGTAAGTGGCA 1641
QY 502 agcgtgtgtcgaatcaatctgggcgttaagaaggtgctgagcgaattgttaagtcagttgtg 561
DB 1640 ACCGTTGTCGGAATTAATTTGGCGTAAGACCGCGCGCGGCTTTCTTAAGTCTGATGTG 1581
QY 562 aaaaactgcggagctcaaccgcgtgcctcctgaactcaactcaagtcctgagatttggaagag 621
DB 1580 AAAGCCCCCGCTCAACCGCGAGAGGTCAATTGAAACTGGGAACCTTGAGTGCAGAAAG 1521
QY 622 gcaagtggaattccagtgtagcgttgaaatgctgtagataltctggagaagaaccagttgac 681
DB 1520 GAGAGTGAATTCACAGTGTACGCGTGAATATCGTATGATGAGGAACACCACTGCGC 1461
QY 682 gaagcgactctgtcgtctaaactgagcgtgagcagaagaagcgtgtgtagtaaacgg 741
DB 1460 GAAGCGACTCTCTGTCTGTATGACGCTGAGGCGCGAAAGCGGTGGGAGCGCAACAG 1401
```

```

QY 742 atagatacccccgaatccacgccctaaacgtgtctaccagttgttggggttt--a 799
DB 1400 ATTAGTATACCTCGTGTAGTCCAGCCGCTAAACGATAGTCTTAAGTGTAGAGGTTTCCG 1341
QY 800 accctcgaatgaacacctaacggaatgaagtagacgcctcgtggactatgtcgcgaagat 859
DB 1340 CCTTTAGTGTGACGACCAACGATTAAGCACTCCGCTGGGAGATGAGTGTAAATTCGA 1281
QY 860 gaaactcaaggaatctgaaggggtccgcacaagcgggtggaacatgtgttaattcat 919
DB 1280 GAACCTCAAGGAATTTGAGGGGGGCGCCGACAAAGCGGTGACCATGTGTTAAATTCGA 1221
QY 920 gatccccaaaacacactcacttggttgatcat-gaatcgaatcatgagatatata 978
DB 1220 GCAACGCGAAGAACTTACCAAGTCTTGACATCTCTGTACAAACCTTAGAGATAGGGCTTC 1161
QY 979 gccctc--ggcagattccaaagtcgtcatgtgtgtcgtcagctcgtcgtgtagatgt 1036
DB 1160 CCTTCGGGGGAGAGTACAGAGTGTGATGTTGCTCAAGTGTGTCGTGATGATGT 1101
QY 1037 tgggttaagtcccgcaagcgaagcgaacccct-ctgtagtgtgtcacttaagttggca 1095
DB 1100 TGGGTTAACTCCCGCAACGAGCGCAACCTTGATCTTACCTGCCACATTCAGTTGGCA 1041
QY 1096 ctggtacgaactgcgcgtgacaaacaggaagcgggagctgaactgaatccatcgt 1155
DB 1040 CTCTAAGTGTACTGCCGGGTGACAAACCGAGAGAGTGGGGATGACGTAAATCATCATG 981
QY 1156 gccctatgtccagggccacacagcgtgtactaatgtgcccatacagaaggttcgccaactcg 1215
DB 980 CCCCTTATGACTGGGCTACACACGCTATCAATATGGGCGAAGAGGCGACGAAAGCCG 921
QY 1216 caagaggaggaactctctaaagtcgtcccaagttcggatgtgggtctgcaacttgacc 1275
DB 920 CGAGGCTAAGCCAAATCCCAAAATCTGTTCTCAAGTGGATGCACTGTGCAACTGACT 861
QY 1276 ccatgtagtcgaatcgcgtgaatctgcgatacgaatgcgcggtgaatgaatcgcgg 1335
DB 860 GCGTAGAGCTGGAATCGTATGATGCGGATCAGCATGCCGCGGTGATACGTTCCCG 801
QY 1336 acctgttacacacgcgcgtccacacacactgagtggggagcaccggaatggtcttgtt 1395
DB 800 GCGTTTACACACCGCCGTCACACAGAGTTTGTAAACACCGAAGTCG-GTGAGGT 742
QY 1396 aaccgttaaggaagacgaactactaaggtgaactcgttaaaggggtggaagtcgttaacaag 1455
DB 741 AACCTTTTGAGACGACGCCCGCAAGGTGGGACAGATGATTGGGTAAAGTGTAAACAG 682
QY 1456 taac 1459
DB 681 TAGC 678

RESULT 9
US-09-745-476-1
; Sequence 1, Application US/09745476
; GENERAL INFORMATION:
; APPLICANT: CANON INC.
; TITLE OF INVENTION: Preparation of Poly-hydroxyalkanoic Acid
; FILE REFERENCE: 4351008
; CURRENT APPLICATION NUMBER: US/09/745,476
; CURRENT FILING DATE: 2000-12-26
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: Microsoft Word
; SEQ ID NO: 1
; LENGTH: 1501
; TYPE: DNA
; ORGANISM: Pseudomonas jessenii P161 ; FERM P-17445
US-09-745-476-1

Query Match          54.4%; Score 803.2; DB 29; Length 1501;
```

Best Local Similarity 75.9%; Pred. No. 1.2e-247;
Matches 1121; Conservative 0; Mismatches 318; Indels 37; Gaps 9;

QY 18 aagcctggcgccgctttaaactgaagtcgagcgg-----gtagcata 65
DB 3 aagcctggcgccgagcttaacacatgcagcgagcgagcgagcgttgcctgaa 62
QY 66 cctagcggcgacgagtgagtaaacatgtgtaactctccagctcgagtgagatacttc 125
DB 63 ttcagcgcgagcgagtgagtaactgctaggaatctcgctgtagtgagggaacaagctc 122
QY 126 cgaaggaagaagctaatccgagtagtccgtgtgatacaagaattgtatagtaagatt 185
DB 123 cgaaggaagagctaatccgacata--cgctctacggaggaagaagcggagccttcggcc 180
QY 186 tatgtcttgagatgagcccgccgagcattagctagctgtgaggaatgagctacaag 245
DB 181 ttcgctatcaagatgagctcagatgcgagattagctagctgtgaggaatgagctacaag 240
QY 246 gcgacgatacgtacgagccgctgagaggtgtccggccacaatggaactgagacagctcc 305
DB 241 gcgagagctcgttaactgtctgagagagatgatacagctggaactgagacagctcc 300
QY 306 atactctacggagcgagcagatgaagaatctgtctcaatgagggaaccctgaagcagc 365
DB 301 agactctacggagcgagcagatgagggaatattgacaatgagggaaccctgatacagc 360
QY 366 gacgcgcgtggaacgaggaaggtcttcgagttgtaagttacatagcagaagaaataag 425
DB 361 cagcgcggtgtgtgagaagaaggtcttcgagttgtaagttacatagcagaagaaag 420
QY 426 cagcaaa-----tgtatgatagttaccctcta--aagcaccggtactacag 470
DB 421 cattaaacctaatagctagtgattgttgcgtaacgacagataaagcccgctactactg 480
QY 471 tgcacagagcgccggttaatacgtatgtgtcaagcgtgtgttcggaatactatggcgtaag 530
DB 481 tgcacagagcgccggttaatacagaggtgtcaagcgttaatacgttaactatgaggcgtaag 540
QY 531 ggtgtcgtacgagatctgttaagtcagtgtaaaactgcyggctcaaccctgagctgca 590
DB 541 cgcgcgtaggtgtgtgttaagttgagtgtaaaagcccggtcacaacctgagacgcga 600
QY 591 ctggaactcaagctcgtgaggttggagagggcgaagtgtgaattccagttgtagcggtga 650
DB 601 tcaaaaactcaacagctagagatgagtaaggtgtggaattccgtgtgtagcggtga 660
QY 651 atggtatgatatctgaggaacacacagtgagcgaagcgaactgtgtctcaaaactgag 710
DB 661 atggtatgatatctgaggaacacacacagtgagcgaagcgaactgtgtatctgatac 720
QY 711 ctgagagcagaagaagcgtgtgtagtaaaacgagatagatacccggtatccacgacctaa 770
DB 721 ctgaggtgagcaagaagcgtgtgtagtaaaacagatagataccctgtgtaccacgctaa 780
QY 771 acgtgtctacacagttgtgtgg--ggtttaacctcaagtaacgaactaacgataagt 829
DB 781 acagtgtcaactaagcgtgtgtggaagccttgagctcttagtggtgcagcaataagca 840
QY 830 agaacgccttgaggagctatgtcgcgaagatgaaaactcaagaagatgaacgggttcgca 889
DB 841 tgaacgccttgaggagctatgtcgcgaagatgaaaactcaagaagatgaacgggttcgca 900
QY 890 caagcgttgagagcagtggtgttaattcagatgatacccaaaaactcaactgagcttgac 949
DB 901 caagcgttgagagcagtggtgttaattcgaagcaacgcaagaagccttaacagcgttcgac 960
QY 950 at--ggaactcaatcatgtagagataatgagccttcggcgagatt--cacaggtgtcga 1006
DB 961 atccaacgaacttcacagagatgagtggtgtccttcgggaacacttgagacaggtgtcga 1020
QY 1007 tgggtgtcgcagcagctgtgtgtgagagatgtgtgggtlaagctccgcaacgagcaacccc 1066
DB 1007 tgggtgtcgcagcagctgtgtgtgagagatgtgtgggtlaagctccgcaacgagcaacccc 1066

DB 1021 tgggtgtcgcagcagctgtgtgtgagagatgtgtgggtlaagctccgtaacgagcgaacct 1080
QY 1067 tatcgtatgttgc---tacctaagttgggacatggtacgaactgcgcgtggaacaaag 1123
DB 1081 tgccttaagttacacagacgtatgtgtggcacttaaggaagatgcgcgtggaacaaag 1140
QY 1124 gagggaagcgaggatgaacgtlcaaaatccatcagcttattatgtlcaagggccacagctgc 1183
DB 1141 gagggaaggtgggagtgagacgtcaagttcatcagcttaccagcttcggtgctacacagctgc 1200
QY 1184 tacatgtccgatacaagaaggttcgccaactgcgaaggggagagcttaactctcaaaagctgc 1243
DB 1201 tacatgttcggtacagaaggtgtgcgaagcgcgaaggttgaggttaactcccaaaacga 1260
QY 1244 tccagatcggatgggtgtgtcgaacacgcagcccatgaagtcgaatcgtatgatacgc 1303
DB 1261 tctgagtcggatgcagctcgtcgaactcgagctgtgaaatcgtatgatacgc 1320
QY 1304 ggaatcagatgcgcggttgaatacgttcccgagacttgtatacacacgcgcgtlcaacacac 1363
DB 1321 gaatcagatgtcgcgttgatacgttcccgagcttgtatacacacgcgcgtlcaacacac 1380
QY 1364 ctgaggtggagacacccggaagtggtcttgttaacgttaaggaagacgactacaagt 1423
DB 1381 gggaggtgtgtgtcaccaggaagtag--ctagttcaactctcgaggagcgttatcacaggtg 1439
QY 1424 aaactcgaaggggtgtgaactgtacaaggtac 1459
DB 1440 tgatctatgactggtgtggaagtcgtacaaggtacg 1475

RESULT 10
US-09-748-205-1
; Sequence 1, Application US/09748205
; GENERAL INFORMATION:
; APPLICANT: Canon, Inc.
; TITLE OR INVENTION: Polyhydroxyalkanoate its manufacturing method, and microorgan
; TITLE OR INVENTION: those are used for the method.
; FILE REFERENCE: 4351009
; CURRENT APPLICATION NUMBER: US/09/748, 205
; CURRENT FILING DATE: 2000-12-27
; NUMBER OF SEQ ID NOS: 1
; SEQ ID NO 1
; LENGTH: 1501
; TYPE: DNA
; ORGANISM: Pseudomonas jessenii 161 strain.
US-09-748-205-1

Query Match 54.4%; Score 803.2; DB 29; Length 1501;
Best Local Similarity 75.9%; Pred. No. 1.2e-247;
Matches 1121; Conservative 0; Mismatches 318; Indels 37; Gaps 9;

QY 18 aagcctggcgccgctttaaactgaagtcgagcgg-----gtagcata 65
DB 3 aagcctggcgccgagcttaacacatgcagcgagcgagcgagcgttgcctgaa 62
QY 66 cctagcggcgacgagtgagtaaacatgtgtaactctccagctcgagtgagataacttc 125
DB 63 ttcagcgcgagcggtgagtaactgctaggaatctcgctgtagtgagggaacaagctc 122
QY 126 cgaaggaagaagctaatccgagtagtccgtgtgatacaagaattgtatagtaagatt 185
DB 123 cgaaggaagagctaatccgacata--cgctctacggaggaagaagcggagccttcggcc 180
QY 186 tatgtcttgagatgagcccgccgagcattagctagctgtgaggaatgagctacaag 245
DB 181 ttcgctatcaagatgagctcagatgcgagattagctagctgtgaggaatgagctacaag 240
QY 246 gcgacgatacgtacgagccgctgagaggtgtccggccacaatggaactgagacagctcc 305
DB 241 gcgagagctcgttaactgtctgagagagatgatacagctggaactgagacagctcc 300

OY	306	atctctcttaaggagacagcagatcttaagaaattctgtctccaattggtgggaaaccccttgaaagcgc	365
Db	301	agactcttaacgggagagacagcttgaggaaatattctggcaatttggtcgaaagccttgatccagc	360
OY	366	gaaccgcgtgaaagaaagaggtctccgattctgaattgaagtctacttgagcaagaaataag	425
Db	361	catgcgcgtgtgtgaagaaagagctctcgattctgttaagaacactttaagtcttggtgaggaagg	420
OY	426	caagaa-----tgctgattgtactctcctcta--aaacacggtctaactcgc	470
Db	421	cattaaacttaatacgttaagtgcttttgacgtttaacgcagaaataagcacccggtctaactctg	480
OY	471	tgcacagacccgcgtgtaatacgttaacgttatgcgaacgcttgcttggaatcattcggcgctaaag	530
Db	481	tgcacagacagccgcgttaatacagaaggtgtccaaagcgtttaactggaattactcgttgctaaag	540
OY	531	gtgtcgtgaagccgattctgttaagtcacagctgttgaaatactgggtcgctaaacccgttgscgtcga	590
Db	541	cgcgcgttaagtgcttgtaagtgcttggaatgtgtgaagcccggtcctaacccgttggaactcgtca	600
OY	591	cttgtaaaactcaaaagctcgtgaagctcttggtgaagagacgaatctgcaggttggtatgaa	650
Db	601	tccaacactgaacagcctaaggtatgtgtagaaggtgtgttggaattctcctgttgtaagcgtgaa	660
OY	651	atgcgtatgatatctcgtgaggaacacccacgctgtgcgaagcgcgaactgcgtgcctaaaactcgc	710
Db	661	atgcgtatgatatagaaagagacacccagctgtgcgaagcgcgaactgcgtgcctaaaactcga	720
OY	711	ctggtgcacagaaagccgtgtgtgaatgaacgggatttaataccccggttaactcgaacgctctaa	770
Db	721	ctggtgtgtgcgaagacgtgtgtgtggaagcaaacggatctagaataccccgttgatccaacgctgaa	780
. OY	771	acgttcttctacccaagtttgctgtg--ggtttaaacctctaagtaagaacacctaacgataaagt	829
Db	781	acgattctcaactaagccgttggtgagcccttgagctctctatgtgtgcgacgataacgataaagt	840
OY	830	agaccgccttggtgagactatcgtctgcgaagagtgtgaactccaaggaattgacgtgtgtgcgca	889
Db	841	tgaacgccttggtgagactatcgtctgcgaagagtgtgaactccaatlaattctgacgtgtgtgcgca	900
OY	890	caaacgcgttgagacatgtgtgtttaaattcgcgtatgaataccccaaactccacttggtgtctac	949
Db	901	caaacgcgttgagacatgtgtgtttaaattcgcgaagcgaacgcgaagaactctaccaggtctgtac	960
OY	950	at'ggatctgaatcatgaatgaataataatgaagcccttcgggacaagat--cacagctgtcga	1006
Db	961	atccaatgaactcttcagagagatgattgtgtgtccttcgcgaacatttgagacaggtgtcga	1020
OY	1007	tgtgtgtcgtcgaagctcgtgtcgtcgagatgtttgtgttaagtctccgcgaacgagcgaacccc	1066
Db	1021	tgtgtgtcgtcgaagctcgtgtcgtcgagatgtttgtgttaagtctccgcgaacgagcgaacccc	1080
OY	1067	tactgtatgtgtgc---tacttaagtgttggaacttggtaagaactgcgcgttgaaacacgcg	1123
Db	1081	gtctcttaagttaacagacgttaatgtgtgtgtgcactctcaagagaaactgcgcgttgaaacacgcg	1140
OY	1124	gaggaaagtcgggtgaatgaactgccaatccatcattggtcctttaatgtccaggggcacaacacgtgc	1183
Db	1141	gaggaaagtcgggtgaatgaactgccaatccatcattggtcctttaatgtccaggggcacaacacgtgc	1200
OY	1184	tacaaatgcgcgaatacagaaggtctgcgaactcgcgaacccatggaatctgggatcgtctagttaatgc	1243
Db	1201	tacaaatgtctgtacagaaggtctgtccaaacgcgcgaaggttggaattactccacaacaaaccca	1260
OY	1244	tcccaattctggaattggtgtctgcgaactcgcgaacccatggaatctgggatcgtctagttaatgc	1303
Db	1261	tctgaatctcgggaatcgcgaagcttcgcgaacccgcgaactgcgtgtgaagtcgggaatcgtctagttaatgc	1320
OY	1304	gatatcgaatctcgcgtgtgaatacgttcccggaacctctgatacacgcgcgttcaacacac	1363
Db	1321	gatatcgaatctcgcgtgtgaatacgttcccggaacctctgatacacgcgcgttcaacacacat	1380
OY	1364	ctgagvtggtggaaccccggaagvtgctcttggttaacccgttaaggaagacgaactaagtg	1423

Db 1381 ggagagtggtgtgcaccgaagtag -ctagtctaacttcggagagcgttaccacgctg 1439
 Oy 1424 aaactcgtaaaggggggtgaagtcgtacaaggracc 1459
 Db 1440 tgatcattgactgggggtgaagtcgtaccaaagtacg 1475

```

RESULT 11
US-09-791-592-1
; Sequence 1, Application US/09791592
; GENERAL INFORMATION:
; APPLICANT: Canon Inc.
; TITLE OF INVENTION: Polyhydroxyalkanoate containing 3-hydroxybenzoylalkanoic acid
; FILE REFERENCE: 4396021
; CURRENT APPLICATION NUMBER: US/09/791,592
; CURRENT FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 1
; SEQ ID NO: 1
; LENGTH: 1501
; TYPE: DNA
; ORGANISM: Pseudomonas jessenii 161 strain.
US-09-791-592-1

```

Query Match	54.48;	Score 803.2;	DB 30;	Length 1501;
-------------	--------	--------------	--------	--------------

Best Local Similarity 75.98; Pred. No. 1.2e-247;
Matches 1121; Conservative 0; Mismatches 318; Indels 37; Gaps 9;

QY	18	aagctctgcgcgcgcgtctttaacatgcgaagctcgaacg-----ggtagcaata	65
Db	3	aacgctctgcgcgcgcgtctttaacatgcgaagctcgaacg-----ggtagcaata	62
QY	66	cctagccgcaacgctgtagtatacgcgtgtatctctccgagctctggtgataacttc	121
Db	63	ttcaagcgcgcgcgcgtgtagtatacgcgtgtatctctccgagctctggtgataacgctc	122
QY	126	cgaagaagaagctaacccgcgaatagctcctgtgtgatacacaagaattgtatagtaagaat	18
Db	123	cgaagaagaagctaacccgcgaatagctcctgtgtgatacacaagaagcagggagaccttggtgc	180
QY	186	tatctctggaagataagaccgcgcgcgcgttatagctagttgtgtgaggttaaigtctacgaag	24
Db	181	ttgcgtctatacgaatagaccgcgcgcgcgttatagctagttgtgtgaggttaaigtctacgaag	24
QY	246	gcgaagatctggttagcgcgcgccttgagagagggtgtctgcgcgcacaatgtgaactgagacacgctcc	30
Db	241	gcgaagatctggttagcgcgcgccttgagagagggtgtgtgagggatgtacatctgtgaactgagacacgctcc	30
QY	306	atactcttaacgcgcgcgcgcgcgcgtttaagaatctgtctcaatgtgggggataaacctctgaagcgc	36
Db	301	agaactcttaacgcgcgcgcgcgcgcgtttaagaatctgtctcaatgtgggggataaacctctgaagcgc	36
QY	366	gaacgcgcgttgaaacaaagaagaagctctctcgagattgtataaattcaatltagagcagcaaaaataag	42
Db	361	catgcgcgcgtgtgtgaaagaaggtctctcgagattgtataaagcactttaagtgtggaggaaggg	42
QY	426	cagca-----tgtatgatatgtgtactcgtcta--aagcacgcgctaaactacg	47
Db	421	catatacctaatacgtctgattgtcttgacgctctacacgaataaagcacgcgctaaactacg	48
QY	471	tgcgcgcgcgcgcgttgataagctatggtgtgcaagcgtctgttcggaatactatgtggcgttaag	53
Db	481	tgcgcgcgcgcgcgttgataatacgaaggtgtgcaagcgtctaaatcgtgaattctctggcgttaag	54
QY	531	gggtgcgtagcgcgaattgttgaagctcaggtgtgtgaaaacgtgcgcgcctcaaacgcgttgycgtca	59
Db	541	cgcgcgtgaggtgttctgttaagcttgagatgtgaagcccggtctcaaccttggaactca	60
QY	591	cttgaaactatacgaagctctggaagcttggtggagagcgaagtgaattccaggtgtgacggtgaa	65
Db	601	ttcaaacatgcgaagctctgagataaggtgaggtgagagaggttggtgaattctctctgttagcgtgaa	66

Db 961 atccaatgaacttccacagatgtagtgctccttcgggaacatlvagacagtgctgca 1020
Qy 1007 tgggtgtcgtacagctcggtctgtagatgttggttaagtcccgcaagacgcaaccc 1066
Db 1021 tggctgtccgaagctcggtctgtagatgttggttaagtcccgcaagacgcaaccc 1080
Qy 1067 tatcgtatgttgc---taccttaagtctggtacgtctgtaaacctgcggtgacaaccc 1123
Db 1081 tgcctctagttaaccagacgtatgttggtgacacttaaggagacgtccggtgacaaccc 1140
Qy 1124 gaggaagcggtgtagatgacgtcaaatccatcgtccttaagtccagggcacaacgtgc 1183
Db 1141 gaggaaggttggtgtagatgacgtcaaatccatcgtccttaagtccagggcacaacgtgc 1200
Qy 1184 tacatgtgcccatacagaggtctgcgaacacgtcgaagggaggtacatctctaaagtgcg 1243
Db 1201 tacatgtgtcgtatacagaggtctgcgaagccggaaggtgaggtacatctccacaacccga 1260
Qy 1244 tcccaatctggaattgggtctgtcgtcaactcgaaccacatagaagtcgtatgaatcgc 1303
Db 1261 tctgtatgcgtacgtcgtacgtcgtacgtcgtacgtcgtacgtcgtacgtcgtacgtcgc 1320
Qy 1304 ggtatcagacgtccggtgtagatgattcccgagacctgtacaacccggtcacacac 1363
Db 1321 gaataagaatgtcgtgtagatgattcccgagacctgtacaacccggtcacacacat 1380
Qy 1364 ctgaattggggagccacgaagtggtctgttaaccgttaaggagagacagactactaagtg 1423
Db 1381 gggagtggtgtgacacgaagtag-ctagtccaacttcggagagagacggttacacgggtg 1439
Qy 1424 aaactcgtaaagggtgtagatgattcgttaacaggtagc 1459
Db 1440 tgaatcactgactgggtgtagatgattcgttaacaggtagc 1475

RESULT 13
US-09-821-016-5
: Sequence 5, Application US/09821016
: GENERAL INFORMATION:
: APPLICANT: CANON INC.
: TITLE OF INVENTION: Polynidroxalkanoate Synthase and Gene Encoding the Same Enzyme
: FILE REFERENCE: 4051021
: CURRENT APPLICATION NUMBER: US/09/821.016
: CURRENT FILING DATE: 2001-03-30
: NUMBER OF SEQ ID NOS: 11
: SOFTWARE: Microsoft Word
: SEQ ID NO 5
: LENGTH: 1501
: TYPE: DNA
: ORGANISM: Pseudomonas jessenii P161 ; BP-7376
: FEATURE:
: US-09-821-016-5

Query Match 54.4%; Score 803.2; DB 31; Length 1501;
Best Local Similarity 75.2%; Pred.No. 1.2e-24;
Matches 1121; Conservative 0; Mismatches 318; Indels 37; Gaps 9;

Qy 246 ggcagatccgttagccggccctgagaggggtgtccggccacaatgtgaactgagacaggtcc 305
Db 241 ggcagatccgttagactgtgtctgagagagatgatacgtacactgtgaactgagacaggtcc 300
Qy 306 atactctacaggagagcaglttaagatctgtctcaatgtggtgggaaacccctgaagcagc 365
Db 301 agactctacaggagagcagltggtgagatattggaacatgtggcgaagagctgtccagc 360
Qy 366 gacgcgcgtgaaagaaagggtctcgtgattgttaagtcttaagcagagaaaaatag 425
Db 361 catgcccgtgtgtgaaagaaagggtctcgtgattgttaagcacttaagtctggaggaagg 420
Qy 426 caagca-----tgtatgattgtgacgtcgtccta--aaagccggttaactgcg 470
Db 421 catcaacttaatacgttgaattgttctgacgttaccagcagataagacacggttaactctg 480
Qy 471 tgcacagccgcgtgtaatacgtatgtgtcaagcggtgtctggaatcaatgtggcgttaag 530
Db 481 tgcacagccgcgtgtaatacgtatgtgtcaagcggttcaatcgttaactgtgtggttaag 540
Qy 531 ggtgtgtagcggatgtttagatcaggtgtgaaactcgtgggtctaaacgtgtgcctgtca 590
Db 541 cgcgcgttaggtgtgtttagatcaggtgtgaaagcccggtctaaactgtggaactgtca 600
Qy 591 ctggaactcaacagctgtgaggttttggaagcgaagtgtgaattccaggtgtgagcgtgaa 650
Db 601 ttcaaaactgacaagctgtagatgtagtagaggtgtgtggaattccctgtgtagcgtgtga 660
Qy 651 atgcgtatgatalcttgagagaaacacagctgtgcgaagcgactgtcgtgtcctaaactgtgc 710
Db 661 atgcgtatgatalaggaagaaacacagctgtgcgaagcgactgtcgtgtatctatgtgca 720
Qy 711 ctgagagccggaagcggtgtgtatgaacgggtattgaatcccggtatccacgctctaa 770
Db 721 ctgaaggtgcgaagcggtgtgtgtgaacagagattgaatccctgtgtatccacgctctaa 780
Qy 771 aggtgtctacaggt 829
Db 781 agcatgttaactaactgacgt 840
Qy 830 agaccgccttggtggtatctgtcgtcaagagtgaaactcaaaagaaatgtgacgggtgtcga 889
Db 841 tgaaccgtctggt 900
Qy 890 caagcggtgtgagcaggt 949
Db 901 caagcggtgtgagcaggt 960
Qy 950 at-ggactgaaatcgtgagagataatgagccttcggtcagagat--cacaggtgtgtca 1006
Db 961 atccaatgaacttccacagatgtagtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1020
Qy 1007 tgggtgtcgtacagctcggtctgtagatgttggttaagtcccgcaagacgcaaccc 1066
Db 1021 tggctgtccgaagctcggtctgtagatgttggttaagtcccgcaagacgcaaccc 1080
Qy 1067 tatcgtatgttgc---taccttaagtctggtacgtctgtaaacctgcggtgacaaccc 1123
Db 1081 tgcctctagttaaccagacgtatgttggtgacacttaaggagacgtccggtgacaaccc 1140
Qy 1124 gaggaagcggtgtagatgacgtcaaatccatcgtccttaagtccagggcacaacgtgc 1183
Db 1141 gaggaaggttggtgtagatgacgtcaaatccatcgtccttaagtccagggcacaacgtgc 1200
Qy 1184 tacatgtgcccatacagaggtctgcgaacacgtcgaagggaggtacatctctaaagtgcg 1243
Db 1201 tacatgtgtcgtatacagaggtctgcgaagccggaaggtgaggtacatctccacaacccga 1260
Qy 1244 tcccaatctggaattgggtctgtcgtcaactcgaaccacatagaagtcgtatgaatcgc 1303
Db 1261 tctgtatgcgtacgtcgtacgtcgtacgtcgtacgtcgtacgtcgtacgtcgtacgtcgc 1320

Db 4063 GTAGC 4059

RESULT 15

US-08-520-946-160

Sequence 160, Application US/08520946

; GENERAL INFORMATION:
: APPLICANT: BROW.

APPLICANT: LYAMICHEV, VICTOR

APPLICANT: OLIVE, DAVID M.

TITLE OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF

; TITLE OF INVENTION: PATHOGENS

NUMBER OF SEQUENCES: 160

CORRESPONDENCE ADDRESS:

ADDRESSEE: MEDLEN & CARROLL
STREET: 330 MONTGOMERY STREET SUITE 3300

STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO

CITY: SAN FRANCISCO
STATE: CALIFORNIA

STATE: CALIFORNIA
COUNTRY: UNITED STATES

ZIP: 94104

COMPUTER READ ;

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS

```

;
; OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0.0
;

```

```

; SOFTWARE:  patcellm Release #1.0, version #1.23
;
CURRENT APPLICATION DATA:
;

```

APPLICATION NUMBER: US/08/520,946

FILING DATE: 06/08/2010

CLASSIFICATION

ATTORNEY/AGENT INFORMATION
NAME: CARROLL DEPTNAME: CARROLL, PETER G.
REGISTRATION NUMBER: 32

REGISTRATION NUMBER: 32,837

TELECOMMUNICATION INFORMATION:
REFERENCE/DOCUMENT NUMBER: FOKS-01/58

TELEPHONE: (415) 705-8410

TELEFAX: (415) 397-8338

; INFORMATION FOR SEQ ID NO: 160:

SEQUENCE CHARACTERISTICS:

LENGTH: 1555 base pairs
TYPE: nucleic acid

TYPE: nucleic acid
STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (ge

US-08-520-946-160

Quorum Watch EA 39.

Query Match
Best Local Stmt

Best Local Similarity	75.3%;	Pred. NO. 4.2e-24/;
Matches 1134:	Conservative	0: Mismatches 324:
		Indels 48:

malices 1334; conservative 0; mismatches 324; inlets 46; gaps 3;

Oy	1	gaatagctcaagacacaaacacgtctggcgcgcgtctttaaactgcaatgcgcgcgcggttag	60
Db	16	GATCTGGCTAGGATGTAACGCTGGCGCGTCCCTAATACATGCAAGTCGAGCGACAGGA	75
Oy	61	caatacc-----tagcggcgacgcggtgaglaaacacgttg-taatcttc	100
Db	76	CGAAGAGCTTCCTCTCTGATGTAGCGCGCGAGCGGTGATTAACCTGGATTAACCTAC	130
Oy	104	ctccgcagctctggagaactattccgaaagaagcgaactaacgcggaatagtcctgttgatca	160
Db	136	CTATAAGAGCTGGGTAACCTTCGGGAAACCGGAGCTAATACCGGATAAI-ATTTTGAAACG	190
Oy	164	caagaattgataagttaaga-----ttatctgcttgagagatgacccgcgcgcgaltta	210
Db	195	CATGTTCTCAAAAAGGAAAGAGACGGCTTCGTCTCACTTAATATGATGATCCGGCGTCATTA	250
Oy	217	gctagcttgtagtagttaatctgctcacaagaagcgacaactcgtctagtcgcgcgtctagaagtggt	270
Db	255	GCTAGTGGTAAAGGTAACGGCTTACCAAGGACGACGATACGTAGCGACACTAGAGGTGA	310
Oy	277	ccgcgcacaatcggaaactgagacacagcttcatactctcctacgcygagcgagcttaagaatc	330
Db	315	TTCGGCCACACAGTGAACCTAGACACGGTTCACACTCTCACTGGGAGGACACACTAGGAATC	370

OY	337	ttctgcatacgaagggaaccccttcgaagcagcagccgcgtgaaacgaagaagccttcgcat	396
Db	375	TTCCGCATG6GCGCAAAAGCCTTACACGAGCAACGCCCGAGTGAATGAAGGCTTCGGAT	434
OY	397	tgtaaaagttcaatcctcgaaggaataaa-----taagcagcaaatgltgatgatgt	443
Db	435	CGTAAAACTCGTATTATGAGGAAGAACATATGCTGTAAGTAACCTGTGCAACTCTTGAAGCT	494
OY	444	acct--gocataaagcacccgcgtactaacgtgtccagcagccgcgttaatacgtatgtgca	501
Db	495	ACCTAATCAGAAACCCACGGCTAACTACGTACGCGACGACCGCGGTAAATACGTAGGTGSCA	554
OY	502	agcgttgttcggaaatcaatctggcgttaaaagcgtgcgaagcggatctgtaaagtcaggtgtg	561
Db	555	AGCGTATTCGGGAATTTATGCGGCTAAAGCGCCCGTAGCGGCTTTTATTAAGTCTGATGTG	614
OY	562	aaaactcgcgcgtcaaacccgttgccgcgtcaccttbaaactcaaaagtcctggagtttggaaag	621
Db	615	AAAGCCCAAGGCTCAACCGTGGAGGGGTCATGSAAACTGGAAACTGTGATGTCAGCAAGACG	674
OY	622	gcaagtggaaatcccaagtgtaagcgtgtgaaatctcgtatagataatcttgaagaaacaagtgcc	681
Db	675	GAAAGTGAATTTCCATCTGTGAAGCGGTGAATATGCGAGAAATATGAGGAMACCAGTGGC	734
OY	682	gaaagcagactctgcctcgtgttcaaaaactaagcttggcgaagaagacgtctgggttaatacggg	741
Db	735	GAAAGCGACTTTTCTGTCGTCTGTAACGCTGTGATGTGCGAAGACGCTGGGGATCAACACAG	794
OY	742	atagataccccgcgttaatccacgcgcctaaacgttgtctacacagttgttgggggttt--a	799
Db	795	ATTGATATACCTTGGTAGTCCACGCGGTAACGATGAGTGTCTAAGTGTATAGGGGGTTTCCG	854
OY	800	acctctagtaacgaaacttaacaggaatgaagtgaagccgcctcggggagctatgctgcgaaggt	859
Db	855	CCCCCTATGCTGCGAGGTAAACGATTAAAGCACTCCCGCTGGGAGTACACCGCAAGGTT	914
OY	860	gaaactcaaaagaaattacaggggggtccgcgaacagcgttggaacatgtgtttaatctgat	919
Db	915	GAACCTCAAAAGATTTACGCGGACCCGCAACAGCGGTGGACGATGTGTTAATTTCGAA	974
OY	920	gatacccccaaaacactcaactcgtggcgttgacatgatalcgtaatc---atgtagagatat	974
Db	975	GCAACGCGCAAAACCTTACCCAAATCTTGACATCTGTTGCAACCTGACGATGACGCTT	1034
OY	975	atgagacctcgggcgaagtttcaacagcgtgtcgtcatgtgttcgtcaagctcgtgtgcggaat	1034
Db	1035	CCCCCTCGGGGGAACAAAGTACACAGGTGTGTCATGTTGTCGTAAGCTCTGTGCGTGAAT	1094
OY	1035	gttgggttlaaagttccgcgaacagcagcgaaccccatcgtatgttgcta--ccttaaaattgg	1093
Db	1095	GTTGGGTTAANTCCCGACAGAGGCGCAACCTTAACCTTAGTTGCCATCATTTAATTGGG	1154
OY	1094	cactgtgtacgaactcgcggtgtgacnaacccgaggaagcggggatgaaocgtacaatcccta	1153
Db	1155	CACCTTAAGTTGACTGCGCGGTGACAAACCGAGGAAGGTGGGGATGACCTCAAAATCATCA	1214
OY	1154	tgaccttaatgttccagggccaacacagttgtataatgycogtataaggaagcgcgcgaact	1213
Db	1215	TGCCCCATTATATTGGGCTACACACTGCTGTAAATGCAACATACAAAGGCGACGGAAC	1274
OY	1214	cgcgaagagggcctaactctcctaaagtcggttcccaagttcggatctgggtgtctcgaactcga	1273
Db	1275	CGCAGAGTCACAGCAAAATCCCATTAAGTGTTCCTCAATTGGGANTTGAATGCTCACAAC	1334
OY	1274	ccccatgaatcggaaatcgcgtagtagtaatcgcgcgatcagcatgcgcggtgtaaatagttccc	1333
Db	1335	CTACATGAAGCTGGAATTCGTAGTAATCGTAGTAGCATGCTACAGCTGAATACCTTCCC	1394
OY	1334	ggaactgttaaacgcgcgcgttcaacaccactgaagtcggggagcaaccggaagcgtgcttg	1393
Db	1395	GGATTATGTATACACCCGCCGTCACCAACGAAAGATTGTGAACACCAGCAAGCCGGTGGAG	1454

Mon Dec 17 07:48:11 2001

us-09-380-826a-1.rtf

Page 16

	1453	1454	1455	1514
OY	1394	1454	1455	1514
Db	1455	1515	1520	

Search completed: December 15, 2001, 05:42:50
Job time: 15815 sec

Db 191 aaagcgggggacgcg--caagaccctgcgcgcgattgtagcgcgattgctccgattagctagt 24

QY 223 tggtagagtaatgcctcacccaagycgacgacatcgtagaccggcctagagaggtgtccggcc 28

QY	283	aababgaacacgagacacggtccatctactcttaagggagcagcagcttaagaactctgtcc	342
Db	309	acactctggaactgaaacacggtccagactctctaaacgtgagacagacagctgaggatatttgc	368
QY	343	aattggggaaaccccttgagacgagacgacgcgcgtgaaacgaaagagctctcgattgataa	402
Db	369	aattggcgcaagacctgtatccaagcattgcgcgcgttgaggtagaagaagaccttcgggtttataa	428
QY	403	gttcattagagcaggaataaataagcag-----caattgtatgtatgtactgtcc	449
Db	429	gacctcttcttcggaaagaataaagcactgtgataatacttcggttctgtcagcgtatccgga	488
QY	450	cta--aaagcacgcgtctaactacgctgtccacgacgcgcgttaatacgtatgvtgacaagcgt	507
Db	489	agaaataagacacgcgtctaactctgtccacgacgcgcgttaatacgaaggtgtcaagcgtt	548
QY	508	gttcggagatacttgagggttaaagggctgctgtagccggaatttgttaactcaaggttgtaaact	567
Db	549	actcgggaattactctggggcttaaaagcgtgctgaagctggttcgttaagctctgtatgtgaaacc	608
QY	568	gcgggtctcaaacccgttgccctgtgcacttgtaaactcaaaagctcttgagatttggaagaggaagt	627
Db	609	ctgggtctcaaaccttggaattgtcatctgtatctgtgcgcgtctagagctgcgttagagatgctg	668
QY	628	ggaattctccaagcttgagagcgggtggaatactgtgatactatcttgagagaaacacacagcttg	687
Db	669	ggaattctcccgctgtagaagcgtgaaatctgctagaaatcctggagagaaacatctctgtgcgaagc	728
QY	688	gactctgtcgtctcaaaactctgacgctctgagacagcaaaagcgttggttagtctaaacgggata	747
Db	729	ggccactctgtgacacgacacttgacactctgagacgaaagcgttgaggagcaaacaggtttaga	788
QY	748	taccgcggtatattccacgccttaaaagctgtctctacacagttgtctggggg--ttttaaccttc	805
Db	789	tacctctgtatgtccacgcctctaaacgaatgcgaactctggaattgtgtgtgcaactatagcacttc	848
-QY	806	agttacgaacacttaacggaattaaagttagacgcgcctcggggactatgtctcgcgaaggttgaaact	865
Db	849	agttctcgaagaacttaacgcggttaagttctgcgcgcctcggggagttacgcgtccgaagacttgaaact	908
QY	866	caaaagggaatttgacgggggtctccgcacaaagcgttgagaaacagctgttcttaattctgaatatacc	925
Db	909	caaaagggaatttgacgggggtcccccgcacaaagcgttgagaaatgattgttcttaattctgaatatacc	968
QY	926	ccaaaacacttcaactctgggctcttgacatggaatctgaaat--caattagagatatatgaagcct	982
Db	969	cgaagaacacttataccttgccctcttgacatgacgcgaaccttcacagagatgagttgtgtgccttc	1028
QY	983	tcgggcagatattcaacaggtgtctgtacatgtgtctgtctgtctgaagctcgtgtctgtgagatgttggtt	1042
Db	1029	ggggaaacctgtgacaaaggtgtctgtacatgtgtctgtctgtctgaagctcgtgtctgtgagatgttggtt	1088
QY	1043	aaggtcccgcaacagcgcgaacccctatactgtattgttc--taccttaagtctggacacggtg	1099
Db	1089	aaggtcccgcaacagcgcgaacccctctgtccctagtgtgccagacgataatgvtgvgaaactct	1148
QY	1100	tacgaaacctgcgcgttgacaaaacccggaggaagcgggagatgacgttcaaatctctatgtgcct	1159
Db	1149	aagggaagacccgcgcgcgcgaacaaagccggaggaaggttggtgatagtacaaatcattcatgtgcc	1208
QY	1160	ttatgtctcaagggtccacaacagctgtctacaatctggccgataaaggggtctgcacaaacttgacaag	1219
Db	1209	ttatgggtccaaagggtccacaacagctgtctacaatctgggtgaggaacaaagggtctgcaaaactctgcag	1268
QY	1220	agggaagcttaactctctaaanaatctgctgtccagcttcgcgaattgggtctgcgaactgtgacccat	1279
Db	1269	agggtgagcgaattccacagaataaccccatctctcgttccgcgttggtgattctgcgaactctgactccat	1328
QY	1280	gaagctcggaactcgtctagttaactcgtgcgatacagca--tgccgcggttgataacgttctccgggac	1338
Db	1329	gaagctcggaactcgtctagttaactcgtcagatacgaatctgtcgcgttgataacgttctccgggac	1388
QY	1339	ctgtatacacacgcgcgtctacacacacacctgagcttggtggagacacccgaaagtgtctctgttaac	1398

[illegible]

```
|||||
Db 473 cgtgcacgacgcgcgtaatacgtaggtctgcgagcgtctgcggaattacttgcgcaaa 532
QY 529 aggtgtgtgagcgcgatttgaatcaggtctgaaatcgcggcctcaaccgtgacctg 588
Db 533 agagctcgttaggcggtctctgcgcctctctgaaacccgcgaactcaacgcgcgacctg 592
QY 589 cacttgaacatacaagcttgcgttgcgagagcaagtgaaattccgcggtgtagcggtg 648
Db 593 caggcgatacgcgcgacgtctgagctcgcagagggagagtaattccgtgtagcggtg 652
QY 649 aaatgcctagaatactctgagaggaacacacagctgcgaagcgagctctgctgcctcaaaatga 708
Db 653 aaatgcctagaatactcagagaggaacacacagctgcgaagcgagctctgctgcgtagtaactga 712
QY 709 cgtctgagagcagaaagcgttgcgttagtaaacgggattagatacccgctgaatccagccct 768
Db 713 cgtctgagagcagaaagcgttgcgttagtaaacgggattagatacccgctgaatccagccct 772
QY 769 aaagcttgcctacacagctgtgtgcggttttaacc---tcagtaacgagcctaaccgagtt 825
Db 773 aaacgctgtgcgcgcgcagagctgtgtgttccctccacggagctccgtgcgtaagcctaaccgact 832
QY 826 aaatgaacgcgccttgcgagactatgctcgcgaagatgaaactcaaaagaaattgacgcgggctc 885
Db 833 aaagccccgccttgcgagactatgctcgcgaagatgaaactcaaaagaaattgacgcgggctc 892
QY 886 cgcacaaagcgcgttgcgagactatgctcgcgaagatgaaactcaaaagaaattgacgcgggctc 945
Db 893 cgcacaaagcgcgttgcgagactatgctcgcgaagatgaaactcaaaagaaattgacgcgggctc 952
QY 946 tgcatacgtatcgaatacgtatgagagat-ataagagccttcgcgagcctaaccagctgtctg 1004
Db 953 tgcatacgtatcgaatacgtatgagagat-ataagagccttcgcgagcctaaccagctgtctg 1012
QY 1005 catgtgtctgcacagctcgtctgcgtgagatgttggttaagctccgcgaacgcgcgaac 1064
Db 1013 catgtgtctgcacagctcgtctgcgtgagatgttggttaagctccgcgaacgcgcgaac 1072
QY 1065 cctatcgtatgttc---taacttaagtctggcactgtgtgaagaactcgcgttgaacaa 1121
Db 1073 cctatcgtatgttc---taacttaagtctggcactgtgtgaagaactcgcgttgaacaa 1132
QY 1122 cggaggaagcgcgttgcgaatacgtatcgaatacgtatcgaatacgtatcgaatacgtatcga 1181
Db 1133 cggaggaagcgcgttgcgaatacgtatcgaatacgtatcgaatacgtatcgaatacgtatcga 1192
QY 1182 gctacaatgcgcgaatacgtatcgaatacgtatcgaatacgtatcgaatacgtatcgaatacgt 1241
Db 1193 gctacaatgcgcgaatacgtatcgaatacgtatcgaatacgtatcgaatacgtatcgaatacgt 1252
QY 1242 ggtccacgttgcgtatgtgcgtcgaatacgtatcgaatacgtatcgaatacgtatcgaatacgt 1301
Db 1253 ggtccacgttgcgtatgtgcgtcgaatacgtatcgaatacgtatcgaatacgtatcgaatacgt 1312
QY 1302 ggcgatacgc-atgcgcgcgtgcgaatacgtatcgaatacgtatcgaatacgtatcgaatacgt 1360
Db 1313 ggcgatacgcgcgcgtgcgaatacgtatcgaatacgtatcgaatacgtatcgaatacgtatcga 1372
QY 1361 caacttgaatgcgcgaatacgtatcgaatacgtatcgaatacgtatcgaatacgtatcgaatacgt 1418
Db 1373 catgaagaatgcgcgaatacgtatcgaatacgtatcgaatacgtatcgaatacgtatcgaatacgt 1432
QY 1419 aggtgaaactcgttaagaggggtgaaagtctgtaacaaagtacc 1459
Db 1433 aggtgaaactcgttaagaggggtgaaagtctgtaacaaagtacc 1473
```

RESULT 3
US-09-622-201-2
; Sequence 2, Application US/09622201
; GENERAL INFORMATION:
; APPLICANT: MATSUYAMA, Akino bu et al.

```
;; TITLE OF INVENTION: Novel microorganisms and processes for producing amide  
;; TITLE OF INVENTION: compounds  
;; FILE REFERENCE: 2224-0169P  
;; CURRENT APPLICATION NUMBER: US/09/622,201  
;; CURRENT FILING DATE: 2000-08-14  
;; PRIOR APPLICATION NUMBER: JP P1998-356229  
;; PRIOR FILING DATE: 1998-12-15  
;; NUMBER OF SEQ ID NOS: 2  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO: 2  
;; LENGTH: 1480  
;; TYPE: DNA  
;; ORGANISM: Unknown Organism  
;; FEATURE:  
;; OTHER INFORMATION: Description of Unknown Organism:16S rRNA gene of  
;; OTHER INFORMATION: Rhodococcus sp. Amb (FERM BP-6595)  
US-09-622-201-2
```

Query Match 52.9%; Score 780.6; DB 6; Length 1480;
Best Local Similarity 75.2%; Pred. No. 2e-259;
Matches 1114; Conservative 0; Mismatches 334; Indels 33; Gaps 10;

```
QY 4 catggtcagaactcaacgccttgcgcgcgtcttaaacatgaaagctgcgagcgaggtaagcaa 63  
Db 1 cctggtcagagcagaacgccttgcgcgcgtcttaaacatgaaagctgcgagcgagtaagcgc 60  
QY 64 tacc-----tagcgcgcgaacgcgttgcgttaaacgt-gtaactctctccag 110  
Db 61 agcttgcctggttgagattgtgtgcagcgggtgagtaaacgctggtgtgactgccttcgac 120  
QY 111 tctggagataactcttcgcgaagaaagcctaatacgcgatacgtctgttgatcaacaagatt 170  
Db 121 ttcggagataacccgcgggaactgtgttcaatacgcgatacgtatccatgagct-gcatgtct 179  
QY 171 tgaatgaataagatttatgtcttgcgagatgagccgcgcgcgagatagctagtgtgagag 230  
Db 180 cgtgtgtgaaagttactgtgtgcagagtgagccgcgcgctcatcagcttgtgtgaggg 239  
QY 231 taatgtccacccaagcgcagatcgtgtgagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 290  
Db 240 taatgtccacccaagcgcagatcgtgtgagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 299  
QY 291 actggaacacgctccatacctcctaagagagcagcagtaagaaactcgtccataatgaggg 350  
Db 300 actggaacacgctccatacctcctaagagagcagcagtaagaaactcgtccataatgagggcg 359  
QY 351 aaacctgaagcagcagcgcgcgttgcgaagaaagagtgcttcogattgttaagttaacta 410  
Db 360 aaagcctgatacagcagcgcgcgttgcgaagaaagagtgcttcogattgttaagttaacta 419  
QY 411 ggcaggaataaataagcagaacatgtgatagttactctccta--aagacgcgcctaacta 468  
Db 420 agcagggcgcgaagcgcgaag-----tgacggtactcttcagaagaagacgcgcgaacta 472  
QY 469 cgtgcgaacacgcgcgcgtgcgaatacgtatgtgcgaacgcttcttcggaatcattgcgcgtga 528  
Db 473 cgtgcgaacacgcgcgcgtgcgaatacgtatgtgcgaacgcttcttcggaatcattgcgcgtga 532  
QY 529 aggtgtcgtagcgcgattgttaagtcaagtgctgaaactcgcgcgtcaaccgcgtgacctg 588  
Db 533 agagctcgtagcgcgttctgtcgcgtcgtctgtgaaacccgcgaagctcaacgcgcgcgtctg 592  
QY 589 cacttgaacatacaagctctgagatttgcgagagcgaagtgaaattccgcggtgtagcggtg 648  
Db 593 caggcgatacgcgcgacgtctgagctcgcagagggagagtaattccgtgtagcggtg 652  
QY 649 aaatgcctagaatactcgtgagaggaacacacagctgcgaagcgagctctgctgcctcaaaatga 708  
Db 653 aaatgcctagaatactcagagaggaacacacagctgcgaagcgagctctgctgcgtagtaactga 712  
QY 709 cgttgaagcagaaagcgttgcgttagtaaacgggattagatacccgctgaatccagccct 768  
|||||
```



```

QY 838 tggggaactatcgcgaagatgaactcaagaatctgacggggtccgcacacgcgt 897
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 816 tgggaagtaacggccgaagatgaactcaagaatctgacggggtccgcacacgcgt 875
QY 898 ggaagatggttaactcgaatgatacccaaaaacccaactgagctgacatgactc 957
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 876 ggaagatggttaactcgaatgatacccaaaaacccaactgagctgacatgagga 935
QY 958 gaactatgtagat-----atatagccttcgggcagatctcaacggtgtgtgactg 1012
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 936 gctgtactcgaagatctgtatctccccaagggacccctccgcacacgctgtgcattg 995
QY 1013 tctcaactcgtgtcgtgacatgttgggttaagtcccgcaacgacgcaacccctacgt 1072
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 996 tctcaactcgtgtcgtgacatgttgggttaagtcccgcaacgacgcaacccctacgt 1055
QY 1073 atgttc---taacttaagtgtggcactgtgtaacgaactcgcgtgacaaacggagaa 1129
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1056 tagtgcacgacgcttgggtgtggtcactctagaagaaactgcgtgacaaacggagaa 1115
QY 1130 ggcggggaatgacgtcaactcctcatgtgcttattgtccaggccacacagtgtaaat 1189
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1116 ggtggggaatgacgtcaactcctcatgtgcttattgtccaggccacacagtgtaaat 1175
QY 1190 ggcggggaatgacgtcgcgaactcgcgaagagggagtaactctcaaaagtcgtccag 1249
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1176 ggcggggaatgacgtcgcgaactcgcgaagagggagtaactctcaaaagtcgtccag 1224
QY 1250 ttcgatttgggttgcgaactcgcgaaccccaatgaagtcggaatcgtatgtaactgcgga 1309
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1235 ttcgatttgcactcgcgaactcgcgaaccccaatgaagtcggaatcgtatgtaactgcgga 1294
QY 1310 gcatgcgcggtgtaatacgttcccggaacttgtacacacgcgcgttaccacacactg 1359
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1295 gcatgcgcggtgtaatacgttcccggaacttgtacacacgcgcgttaccacacactg 1354
QY 1370 ggggaagacccgaagtggtcttgttgaacgtaagaggaactaagaagtgaactc 1429
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1355 ggtgttgaacngaaacg-gtgaagcgaacgcaagagcgaacgcaacacagtgctgta 1413
QY 1430 gtaaaagggtgaaagtcgtaacaag 1454
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1414 gcgactgggtgaaagtcgtaacaag 1438

```

```

RESULT 5
US-09-897-516-423
; Sequence 423, Application US/09897516
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Huesling, Joseph E.
; APPLICANT: Krasomil-Osterfeld, Karina C.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Slater, Steven C.
; APPLICANT: Spolidorov, Sergei

```

```

; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51847)B
; CURRENT APPLICATION NUMBER: US/09/897,516
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215, 161
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 8409
; SEQ ID NO 423
; LENGTH: 83080
; TYPE: DNA
; ORGANISM: Xenorhabdus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (38740)..(39948)
US-09-897-516-423

```

```

Query Match      51.7%; Score 763.4; DB 6; Length 83080;
Best Local Similarity 73.4%; Pred. No. 1,9e-252;
Matches 1100; Conservative 0; Mismatches 356; Indels 43; Gaps 8;

QY 1 gatactgctcaagaactaagcgttgcgcgcgtcttaacatgcaagtcagcgggtag 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 616 gatcatgctcaagaactaagcgttgcgcgcgtcttaacatgcaagtcagcgggtag 675
QY 61 caatac-----ctagcgcgcgaacgggtgagtaaacgtgttaactc 102
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 676 ggaagcagcttgcgtcttgcgtacgagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 755
QY 103 cctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 162
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 736 ccgcatggaagggtatataccaactggaacggtgtgtataatccgcatacctcaagga-- 793
QY 163 acaagatttgaagtaaaagattatctgtgtgagatgagaccgcgcgcgcgcgcgcgcgcgc 222
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 794 gcaagtgggggaacttcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 853
QY 223 tgttgaagtatactgctcaacgaagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 282
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 854 aggtgggttaacggtctacactgaagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 913
QY 283 acaatggaactggaagacacgcgtctcaactcctacgcgcgcgcgcgcgcgcgcgcgcgcgcgc 342
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 914 acactggaactggaagacacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 973
QY 343 aatgggggaacccctgaagagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 402
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 974 aatgggggaacccctgaagagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1033
QY 403 gttcatagcagga-----aataagcagcaatgtgagatgtactc 447
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1034 gttacttcaagcgggaaggaagcgttgaagtgaactaacttcaacttgaacttgaactcgcgc 1093
QY 448 gccataacacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 507
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1094 agaagaagacacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1153
QY 508 gttcgaatcattggcgcgttaaaaggtgctgaagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 567
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1154 aatcggaattactgtggcgttaaaagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1213
QY 568 ggcgggtcaacccgtgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 627
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1214 ccggccttaacctggaactgacatcctaagactgtgtactagaagtcctcgttagagggggt 1273
QY 628 ggaattcgaagtgtgacgcgtgaatgctgatactcgtgaagacacacagtgcggaaggc 687
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1274 agaattcgaagtgtgacgcgtgaatgctgatactcgtgaagacacacagtgcggaaggc 1333
QY 688 gacttgcgtcctcaaaactgacgcgttgaagcgcgaacgctgtgtgaagtaaacgggattaga 747
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1334 gccccctggagcgaagactgacgcgtcgaagtgcgaaagcgtgtggggaacaacagattaga 1393
QY 748 taacccgttaactcaagccctaaacgttgtctac--cagttgttggggtttaaactccta 806
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1394 taacctgttagtccacgcgcgttaaacgaltgctgaacttgaaggttltgtgaccttgaacgcgtg 1453
QY 807 gtaacgaactcaagcgttaagttagtagcgcgcgttggggactaagcgcgcgaaggttgaactc 866
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1454 ctccggaagcctaaacgcttgaagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1513
QY 867 aaagaattgaacgggggtcgcacaaacggttgaagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 926
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1514 aaatgaattgaacgggggtcgcacaaacggttgaagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1573
QY 927 caaaactcaacccgtgtgacatgatactgaatcat--gtagagatatagaaccttcg 985
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1574 gaagaaccttacctactctgtacatccagcgaaatccttcagagatatagggagtccttcg 1633

```

OY	986	ggcgagatt--caaaagtgctgcgatggttgttcgtcaagctcgtgttcgtctgtagagatggttggtta	1043
Db	1634	ggaacgcctgcgaagcagctgacgtgcacatgcgcgttcgcacagctcgtgttctgtaattcttggtta	1693
OY	1044	agtcgcccaagagcgcacacccctatcgtatgttcg---taacctaaatttgagccattgct	1100
Db	1694	agtcgcccaagcagcgaacctctatcctcttctgttcgcagcagctaatgtgttggaactcaa	1753
OY	1101	acggaactgcgcgtgtgacaaaccgagagacggtggagtgagtcgaatctcctatgacct	1160
Db	1754	ggagagactgcgcgtgtgataaaaccgagagagctgtggagtatagctcaagctcaatagccct	1813
OY	1161	tatgtccagggggcccaacacgtctgtacaaatggccggaatacagaggtctgcgcaactcggcaga	1220
Db	1814	tacgagctagggctatacaacacgtctgtacaaatggcgcagataatacaagaagcagactcgcgaga	1873
OY	1221	ggagagctaatctctaaaaagctcgtgtccacgtctcggattgggggtcgtcgaactcgcagcccatg	1280
Db	1874	gcaagcgcgaagactataaagctctgtctgtagtccggattggaggtctctgcaactcgcactccatg	1933
OY	1281	aagctcggaattcgtctagtagaatactgcgcgtgatacagatgcgcgctgtgataagcttccggacct	1340
Db	1934	aagctcggaattcgtctagtagaatactgcagtcagatgcgcagctgtgataagcttccggacct	1993
OY	1341	gtacacacgcgcctcgcacacacactctaggtggggagaccccggaagtgctctgttataccg	1400
Db	1994	gtacacacgcgcctcgcacacactctaggtggggagaccccggaagtgctctgttataccg	2052
OY	1401	taagaggaacgcgactactaaagtgtaaactcgttaaaagggtgtgaagtctgtatacaaaagtacc	1459
Db	2053	tcggagagggcgtctacacacttctgtatctatgactgactgtgtgaagctcgtatacaaaagtacc	2111

```

RESULT      6
US-09-897-516-3852
- Sequence 3852, Application US/09897516
- GENERAL INFORMATION:
  APPLICANT: Corbin, David R.
  APPLICANT: Hinkle, Barry S.
  APPLICANT: Huesing, Gregory J.
  APPLICANT: Krasomil-Osterlefeld, Karina C.
  APPLICANT: Malvar, Thomas M.
  APPLICANT: Slater, Steven C.
  APPLICANT: Spolidonov, Sergei
  TITLE OR INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
  FILE REFERENCE: 38-21(51847)B
  CURRENT APPLICATION NUMBER: US/09/897,516
  CURRENT FILING DATE: 2001-06-29
  PRIOR APPLICATION NUMBER: US 60/215, 161
  PRIOR FILING DATE: 2000-06-30
  NUMBER OF SEQ ID NOS: 8409
  SEQ ID NO 3852
  LENGTH: 83080
  TYPE: DNA
  ORGANISM: Xenorhabdus sp.
  FEATURE:
  NAME/KEY: CDS
  LOCATION: (1362)..(1466)
  US-09-897-516-3852

```

Query Match	51.7%	Score 763.4	DB 6	Length 83080
Best Local Similarity	73.4%	Pred. No. 1.9e+252		
Matches 1100; Conservative	0	Mismatches 356	Indels 43	Gaps 8

Oy 1 gatcatgtgctcagaacacaaacgtctggtcggtctttaaatactcaatctgaagcggtgag 60
 Db 616 gatcatgtgctcagaatctgaacgtctggtcggtcgaagccttaacacatctgaatctgaagcggtgaaca 675
 Oy 61 caatac-----*--ctacggtcggaagcgtgtgaataacgtgtglaactct 102
 Db 676 ggaagacgaatctgtcgtcttctgtctgaacgagcggtcgagacgtgtgaataatgctctggagatct 735

QY	103	ccccgagctctgggataatacttccgaaaggaaagcctaataccggaLabgctccgtctgatac	162
Db	736	cccgatctggaggggataataaccatctgnaaacccgtctgctaataaccgataccctctaaggac--	793
QY	163	acaagattctgaataagataattattgctctggagatactggccggccggccgagatctaagt	222
Db	744	gcaaaagtgggggacacctccgggctccaagccatacgtgataaaccagaatggaatctaaact	853
QY	223	tggtgagtgtaatgctccaacaaaggcgagatactgtagtaacgctctgagaggggctccggcc	282
Db	854	agtgtagtgtaacgctcacctaaaggcgagatactccctagctcgtctgtaggagatgaccagcc	913
QY	253	acaaatgagacatgaaacacggtccaactctcaacggagagcgacgagttagaatctctgc	342
Db	914	acaaatgagacatgaaacacggtccaacgagacatctcaacggagagcgacgagttagaatctct	973
QY	343	aatgggggaaaccccttgagacagcgaccccgctgaaacgaaagagagctctcgattgtaaa	402
Db	974	aatggggcgaaagcctctgattgcagccatgacgctgctgtatgaaagagccttcgggtgttaa	1033
QY	403	gtccaattagcgaggaaa-----aataagcagcaatgtgatagtgtacct	447
Db	1034	gtaccttaagcggggagaaaggcgctgaaagtgtgataactgcttccagatctgaactccggc	1093
QY	448	gctctaagccacccgctctaactacgtctgcacagagccggcggttaatacgtatgtccaagctt	507
Db	1094	agaagaagaacacccgctctaactccgtctgcacagagccggcggttaatacaggaaggtgtcaagcgtt	1153
QY	508	gtccggaatcatctggcgctgaaagggtgcgtgtagcgagatttgtaagtcaggtgtgaaaact	567
Db	1154	aatcggaattactctggcgctgaaaggcgacgcgcagagcggtccaattaaagttagtgtgaactc	1213
QY	568	gcgggctcaaacccgtgcgcgcacacttgaaactcaagctctggaatttggagaaaggcaagt	627
Db	1214	ccgggctcaaacccgtgcgcgcacacttgaaactcaagctctggaatttggagaaaggcaagt	1273
QY	628	ggaattccaggtctgtagcggtgtaaaatcgttatgatactctgtagaggaacacacccgctgtagaagc	687
Db	1274	agaattccaagctgtagcggtgtaaaatcgttatgatactctgtagaggaataccgctgtagaagc	1333
QY	668	gaattctgcgtctcaaaactctgacgcctgtaggcacagaaagcgctgggttagtaaacgggattga	747
Db	1334	ggccccctggagcgaagatctgacgcctcaagctgtaggcacagaaagcgctgggttagtaaacgggattga	1393
QY	748	taccgccgtaatccacaccccttaaacgttgtctacg-cagttgctctgggggttttaaacctcca	806
Db	1394	taccctgtgtagtccaacgcgttaaacagatgtcgacttgaggggttgttgagccttgagacccgtg	1453
QY	807	gtaacgaaacctaaacgataagtagaagccgctggggagctatgactccgcgcaagagtgaaactc	866
Db	1454	cttcgcggagacttaaacgcgtttaagtcgacgcgccttgggagtagcgtcgcaagattaaaactc	1513
QY	867	aaaaggaattaaacgggggttccgcacaagcggttgtagagatgtgtttaatctcagtagtaacc	926
Db	1514	aaaaggaattaaacgggggttccgcacaagcggttgtagagatgtgtttaatctcagtagtaacc	1573
QY	927	caaaaacctcaacctgggcttgtagatactgatacat-gtagagatatataatgaaaccttcg	985
Db	1574	gaagacacttaaccttaactcttgatcatccaggaatacctctcagaagatgagggagtgctctcg	1633
QY	986	ggcgagatt--cacaggtgcctgcatgtgttctcgaactcgtctgctgtagagatgttggtta	1043
Db	1634	ggaagcgttgagaaacggtgcgtgcatgtgtctgctgcaactcgttgtgtgaaatgttggttta	1693
QY	1044	agtcgccgaaggaacgaacctcatctgtagttgg--taccttaagttggaacccgctgt	1100
Db	1694	agtcgccgaaggaacgaacctcatcttcccttctgttcgacgaacgaatggttggaactcaa	1753
QY	1101	acgaaacctgcggtgacaaaacccgaggaagcggggtagagctcaactctcaatctcgtacct	1160
Db	1754	ggggaggtgcgggtgataaacccggaaggaagtggtgatactcaactcaatctcaatctcgtacct	1813

QY	1161	tatgttccagggtccacatacgtgtgtctataatggtccgatatacagaggtgtctgcgcaactctgcgaaga	1220
Db	1814	tacagataggaggtctacacacgtgtgtctacaatggtgcagatatacaagaagaaggtactctgcgcaga	1873
QY	1221	ggaggtcaactctctcaaaagtcggtctccagttccgatttcggatttgggtgtctcgaactgcagcccatg	1280
Db	1874	gcaagcggaaactcataaagtcctgtctgtatccggatttggagctctgtcgaactctgcattccatg	1933
QY	1281	aagtcggaactcgtctagtaatactgcgcatctcagcatgtccgcgctggatgaacgtctcccggaact	1340
Db	1934	aagtcggaactcgtctagtaatactgcgcatctcagcatgtccgcgctggatgaacgtctcccggaact	1993
QY	1341	gtacacacgcgccgtctacacacactctgaagtgaggaggaaccgcggaagtggtgtctgtataacgg	1400
Db	1994	gtacacacgcgccgtctacacacactctgaagtgagggtgtctgcaaaagaagtaag_gttagcttaacct	2052
QY	1401	taagagagacagactactaaggtgaaaaactctgtaaaaggggtgtagtctgtatacaaggtacc	1459
Db	2053	tcgggaagagcgcttacaacacttctgtattctatctgtactgtgggtgtgaagctgcgtatacaaggtaac	2111

```

RESULT 7
US-09-897-516-3853
; Sequence 3853, Application US/09897516
; GENERAL INFORMATION:
APPLICANT: Corbin, David R.
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Huesing, Joseph E.
APPLICANT: Krasomil-Osterfeld, Karina C.
APPLICANT: Malvar, Thomas M.
APPLICANT: Slater, Steven C.
APPLICANT: Spiridonov, Sergei.
TITLE OR INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
FILE REFERENCE: 38-21(51847)B
CURRENT APPLICATION NUMBER: US/09/897,516
CURRENT FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: US 60/215, 161
PRIOR FILING DATE: 2000-06-30
NUMBER OF SEQ ID NOS: 8409
SEQ ID NO 3853
LENGTH: 83080
TYPE: DNA
ORGANISM: Xenorhabdus sp.
FEATURE:
NAME/KEY: CDS
LOCATION: (1931)..(2044)
US-09-897-516-3853

```

	Query Match	51.7%;	Score 763.4;	DB 6;	Length 83080;
	Best Local Similarity	73.4%;	Pred. No. 1.9e-252;		
	Matches 1100;	Conservative 0;	Mismatches 356;	Indels 43;	Gaps 8;
OY	1 gatcatgctcggaactaacgtctggcgcgcgtctttaacatgaagaagtgcagcggttag	60			
Db	616 gatcatgctcgcatltaacgcgttgcgcgagccctaaccacatgcagcaagtgcacggtlaaca	675			
OY	61 caatc-----ctagcgggaacgggttgataccaacgtttgatctt	102			
Db	676 ggaagcagcttcgtcctttgtctgcacgaaggcggcgagcggttgagattatgtctcgggattcgg	735			
OY	103 cctccgagctcttggaataactttccgaagaagaaagctlaataccgggtatgccgtttgattc	162			
Db	736 ccggttgtaggggtagtaaccactgcgaagaacgggtgattaataccgatgaaccttaagaa--	793			
OY	163 acaaagatttgataaglaaagatttatgtcttgagatgagcccgcgcgcgcgattagattag	222			
Db	794 gaaaagtgsggggaaccttcgcgcctcaacgcgcattcgatgaaccaagattgattagtagt	853			
OY	223 tggvtgaggtaatatgctcaccaagacgaagatcgtgtagccgccttgaaagaggtgtccgcgc	282			
Db	854 aggttggttgaaacgctgcacctagtcgacacatccctcgtctgcttaagaagaatgacacacg	913			

OY	283	aacatggaactgagacaaggttcacatacctccacgagagcgatlaangactctgtcc	342
Db	914	aacatttgactgtagacacacgcccacgactccacgagacagctgggactgtgtac	973
OY	343	aattggggaaaccttgaaagcagcgcgcgttgaaacgaaagagctcttgattgttaa	402
Db	974	aattgggcgacagccttgatgacacagtcgcgcgtgtgtatgaaagaagccttcgggttaaa	1033
OY	403	gttacttagcaggaanaa-----ataagacagaactgtatgatgtact	447
Db	1034	gtactcttcagcgcggagaggaagcgtgtgaagtgtataactgttccagattgacgttaaccgc	1093
OY	448	gctcaaaagcacccggtctaactacgtgtccagacagccgcgtataactgattgtgcagaagcgt	507
Db	1094	agaagaagacacccggtctaactccgtgtccaaacagccgcgcgtataactgaaaggtgtcaagcgtt	1153
OY	508	gttcggatcatcttggcgttaaagggtgcgttgtagccgattctgtaactcaggtgtgaaact	567
Db	1154	aattcggaattacttggcgtgttaaagcagcagcgcgttcaattaaatttgatgtgtaaaccc	1213
OY	568	gcgggcgtcaaccgcttgctgcacttgaaactaacaagcttgagatttgggaagggcaagt	627
Db	1214	ccgggcgtcaacccttggaactgtcacataaagactgtgtactaaagatgctcgtatgaagggggt	1273
OY	628	ggaattccaggtgtgagaggtggaatctcgttagatattctggaaagaacacaggttgccaaagc	687
Db	1274	agaattccacgctgtagcgggtggaatctgcgttagagatgttgtaggaataaccggtgcgcgaagc	1333
OY	688	gactctgctggtctcaaaactgtacgcgtgaggcacagaagcgcgtgggtgtgataaccggataga	747
Db	1334	ggccccctgagcagagactgtacgcctcaggtgcgaagagcgtggggagcaaacagattaga	1393
OY	748	taccgccgtatccacgcgccttaagttgtttac-cagttgttgggggttttaaccctca	806
Db	1394	taccctgtgattccacgcgcgtlaaacgattgtgcacttgaggtgttgcgtctggaagccgcgg	1453
OY	807	gtaacgaacctaaagcgtatlaagtagacccgccttgggagacatctctgcagaagtgtagaac	866
Db	1454	cttcgcgagctaaagcgtttagtgcagccgccttgggagtgagtcgtccgaagattaaactc	1513
OY	867	aaaggaattgacgggggttcgcgcacaagcgggtgagcagatgtgttcaattcgatgtatacc	926
Db	1514	aaatgaatttgacgggggcgcgcacaagcgggttgagagatgtgttcaattcgatgtataacc	1573
OY	927	caaaacctcaaccctgggcgtgacatgatactgtgatcat-gtagagatatatgaagccttcg	985
Db	1574	gaagaaccttaactaactcttgataactcagcaggaatcctcttaagagatgagggaggtgcctcg	1633
OY	986	ggcgcgatt--caacggtgctgcattgtgtctgcagctcgtgtcgttgagatgtgtggtta	1043
Db	1634	ggaacgcgtgagacagcgtgctgcattgtgcctgcgcagctcgttgttgaaatgtcttggtta	1693
OY	1044	agtcgccaaagagacgaacccctatcgtatgtttgc--taccttaagtttgggcacttgt	1100
Db	1694	agtcgccaaagcgcgaacccctatccttctgttgcgcagacagttatggttggaaactcaa	1753
OY	1101	acgaactcgtccgggtgtgacaaaacgcggagagcggggatgacgttcaatctcatgtgcctt	1160
Db	1754	ggggagactcgcgcgtgtgataaacgcggagaggtggggaatgagcttcaagtattcatgtgcctt	1813
OY	1161	tatgtccagggccacaaacgttgtctataacttggccgatacagaaggggtcgcgaactcgcgaaga	1220
Db	1814	tacagatgaggtctacacacgtgtctataacttggcagatacaaaagagaagcgaacctcgcgaaga	1873
OY	1221	ggggcgttaattcttcaaaagcgcgttccaggttgcgatttggggtctgcaaatcgaaacctga	1280
Db	1874	gcaagcgcgaactcaataagctctgtgtgattgcggatcttggaatctgcaactcgaaacctga	1933
OY	1281	aagctcgaactcgtagttaactcgcggtacagcatgtccgcgcgtgataactgcttccgcgaact	1340
Db	1934	aagctcgaactcgtcagtaactcgttagatccagaatgacgtgcaggtgtaataactgtccgcgcgt	1993


```

: GENERAL INFORMATION:
: APPLICANT: Cordin, David R.
: APPLICANT: Goldman, Barry S.
: APPLICANT: Hinkle, Gregory J.
: APPLICANT: Krasomil-Osterfeld, Karina C.
: APPLICANT: Malvar, Thomas M.
: APPLICANT: Slater, Steven C.
: APPLICANT: Spiridonov, Sergei
: TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences and Uses Thereof
: FILE REFERENCE: 38-21(51847)B
: CURRENT APPLICATION NUMBER: US/09/897,516
: PRIOR FILING DATE: 2001-06-29
: PRIOR APPLICATION NUMBER: US 60/215, 161
: PRIOR FILING DATE: 2000-06-30
: NUMBER OF SEQ ID NOS: 8409
: SEQ ID NO 3855
: LENGTH: 83080
: TYPE: DNA
: ORGANISM: Xenorhabdus sp.
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (7060)..(8658)
: US-09-897-516-3855

Query Match      51.7%; Score 763.4; DB 6; Length 83080;
Best Local Similarity 73.4%; Pred. No. 1.9e-252;
Matches 1100; Conservative 0; Mismatches 336; Indels 43; Gaps 8;

QY 1 gatcatgctcaagaactacgtctgacgagcgctcttaacaatgcaagtcgagcgagtag 60
Db 616 gatcatgctcagatctgaacgctgagcgagccctcaacaactgcaagtcgagcgtag 675
QY 61 caatac-----ctagcgagcaagcggtgagtaacaactgtaactt 102
Db 676 ggaagcagcttctgcttctgtctgacgagcggtgagtagtaatgtctcgagatctg 735
QY 103 cctccgagctctggataactctccgaaggaagctaatactacgatactcctgttgatc 162
Db 736 ccgagatggagggataataccacttgaaacggtgcttaatactacgctgacccctaaaga-- 793
QY 163 acaagatttgaatgaataagattatctgtctgagatgacgagccgagcgatagctagt 222
Db 794 gcaagatggggacctctcgagccctcaacgcatcgatgaaccacgagatgagtagtagt 853
QY 223 tggcgagatgaatgctcaacgaagcgagatcgatcgagcgagcgaggggtgtccggcc 282
Db 854 aggtggggttaacggtctcaactcctagcgagatcccttagctgtgtctgagaggaatgac 913
QY 283 acaatgagactgagacagagctcaactcctacgagagcgagcgatgaagaattctgtc 342
Db 914 acacttgagactgagacagagcgagcccgagactcctacgagagcgagcgatgagatattgac 973
QY 343 aatgggggaacccctgaagcagcgagcgagcgatgaacgaagaaggtctcgatgttaa 402
Db 974 aatggcgcaagcctgattcgacgcatcgcggtgtatgaagaagcgcttcggtgttaa 1033
QY 403 gtcatatagcagga-----ataagcagaatgtatgtact 447
Db 1034 gtacttcaaggaggaagcgatgaagtgtaatacgtttcagatgtgacttaccgc 1093
QY 448 gccataagcagcgtaactacgtgacgagcgagcgatgaatgagtcgaagcgt 507
Db 1094 agaagaagcagcgtaactcgtgacgagcgagcgagtgtaataaggaagcgagcagcgt 1153
QY 508 gtccggaatcattggcgtaaaagggtcgtagcgagatttgaagtcagtgtaaaact 567
Db 1154 aatcggaattactggcgtaaaagcgcaagcgagcggtcaatlaagtagatgtgaatcc 1213
QY 568 gcgggctcaaccgctgagcctcaacttgaataaagtcgtgaggttggggaagcaggt 627
Db 1214 ccgggctcaaccctggagactgtacataagactggttgaactagagttcgttagaagggt 1273
```

```

QY 628 ggaattccaggttagcggtgaatatcgtagatatcggaggaaccacgattggcgaagc 687
Db 1274 agaattccactgttagcggtgaatatcgtagatatcggaggaaccacgattggcgaagc 1333
QY 688 gactgtctgctcaaaacttgacgctgagcgacgaagcggtggtgtagtaaacggagtag 747
Db 1334 ggcctccgagcgaagactgagcgcctcaggtgctgaagagcggtggggaacaacaggttag 1393
QY 748 taaccgggttatccagccctcaaaagttgtctac--cagttgttgggggtttaaccctca 806
Db 1394 taacctgttagtccacgctgaagaacgactgtgaggttgttgcttgaagccttgg 1453
QY 807 gtaagaaactcaacgagatgaagtagaacgcttggagactatgtctcgacgaagtgaaact 866
Db 1454 ctcccgagactcaacgagcgttaagatgacccgcttgggagtagcgttcgcaagattaaact 1513
QY 867 aaaggaattgacgggggtcgcacacaagcggtgagagcattgtgttaattcatgataacc 926
Db 1514 aatggaattgacgggggtcgcacacaagcggtgagagcattgtgttaattcatgataacc 1573
QY 927 caaaactcactcgggtctgacatgagatcgaatcat--gtagagatatgagccttcg 985
Db 1574 gaagaaactcactcactcttgcacatccagcgaatcctcaagagatgagaggtgcttcg 1633
QY 986 ggcagatt--cacaagtgctcatgtgtgtctgctcaagctgctgtctgtagatgtgtgta 1043
Db 1634 ggaagcgtgagacaggtgtctgacatggtcgtctcgcgtccgctcgtgtgtgaaatgtgtgta 1693
QY 1044 agtcccgcaacgagcgcaaccctcatcgtatgttcg---tacactaaatttggcagctgt 1100
Db 1694 agtcccgcaacgagcgcaaccctcatccttctgttgcacgacgtaattgttggaaactcaa 1753
QY 1101 acgaactcgcggtgacaaaccggaagcggtgagatgacgttcaaatctcactatgacct 1160
Db 1754 gggagactcgcggtgataaacgcggaaggttgggagactgaactcaatcatgacct 1813
QY 1161 tatgtccagcgccacacacggtctacaaatggtccgatatagaggggtgcgcaactcgcaaga 1220
Db 1814 taaggtatgggtctacacacggtctacaaatggtccgatatagaggggtgcgcaactcgcaaga 1873
QY 1221 gggagactatcttaaaagctggtccagctcgagatttgggggtctgcaactcgaccccatg 1280
Db 1874 gaaagcggaactcataaagatctgtctgtagtccgagtttggagttcgcgaactcgaccccatg 1933
QY 1281 aagtcggaatcgtctagatcgtcgagatcagcatcgcggtggaatcagcttcccgacct 1340
Db 1934 aagtcggaatcgtctagatcgtcgagatcagcatcgcttgaatcagcttcccgacct 1993
QY 1341 gtacacacgcccgtcacacacacccctgaatggggagcaccggaagtgtgttgaacg 1400
Db 1994 gtacacacgcccgtcacacacacccctgaatgggggttgcgaaagaagtag--gtatgctaaact 2052
QY 1401 taaggaagacagactactaaggttgaactcgttaaaagggttgaagtctgataaaggttacc 1459
Db 2053 tcgggagggcgcttaaccacttgttatcatgactgtgggttgaagtcgtaacaaggttaac 2111

RESULT 10
US-09-897-516-3856
: Sequence 3856, Application US/09897516
: GENERAL INFORMATION:
: APPLICANT: Cordin, David R.
: APPLICANT: Goldman, Barry S.
: APPLICANT: Hinkle, Gregory J.
: APPLICANT: Krasomil-Osterfeld, Karina C.
: APPLICANT: Malvar, Thomas M.
: APPLICANT: Slater, Steven C.
: APPLICANT: Spiridonov, Sergei
: TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences and Uses Thereof
: FILE REFERENCE: 38-21(51847)B
: CURRENT APPLICATION NUMBER: US/09/897,516
```

; CURRENT FILING DATE: 2001-06-29
 ; PRIOR APPLICATION NUMBER: US 60/215, 161
 ; PRIOR FILING DATE: 2000-06-30
 ; NUMBER OF SEQ ID NOS: 8409
 ; SEQ ID NO 3856
 ; LENGTH: 83080
 ; TYPE: DNA
 ; ORGANISM: Xenorhabdus sp.
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (8720)..(10027)
 ; US-09-897-516-3856

Query Match 51.7%; Score 763.4; DB 6; Length 83080;
 Best Local Similarity 73.4%; Pred. No. 1,9e-252;
 Matches 1100; Conservative 0; Mismatches 356; Indels 43; Gaps 8;

QY 1 gatcatggtcagaactacgctgagcgagcgtcttaacaatgcaagtcgagcggtag 60
 Db 616 gatcatggtcagaactacgctgagcgagcgtcttaacaatgcaagtcgagcggtag 60
 QY 61 caatac-----ctagcgagcgagcggtgagtaaacgttgtaactct 102
 Db 676 ggaagcagcttctgcttctgctgagcgagcgagcggtgagtaaacgttgtaactct 102
 QY 103 cctcagagcttgagataactcttcgaagaagataaccgagataagtcctgttgatc 162
 Db 736 ccgagtgaggggagataaccacgtggaacggtgagtaataccgagataagtccttgatc 162
 QY 163 acaagattgataagtaagattatctgttgagataagcccgagcgagataagtc 222
 Db 794 gcaaaatgagggagcctcgagcctcagccatcgagatgaaacccagatgagtaagtc 853
 QY 223 tggtagtgaatgctacacacgaagcgagacgagtcgtagccgctgagaggtgtccggcc 282
 Db 854 aggtgggtagaagcgtcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 913
 QY 283 acaatgagactgagacagcgttcatactccacgagcgagcgagcagcagcagcagcagc 342
 Db 914 acacttgagactgagacagcgttcatactccacgagcgagcgagcgagcagcagcagcagc 342
 QY 343 aatgggggaaacccctgaagcagcgagcgagcgagcgagcgagcgagcgagcgagcgagc 402
 Db 974 aatgggggaaacccctgaagcagcgagcgagcgagcgagcgagcgagcgagcgagcgagc 402
 QY 403 gtcatlaagcagga-----aataagcgaatgtgatgtaact 447
 Db 1034 gtacttcaagcgagggagagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagc 1093
 QY 448 gcttaagcagcgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 507
 Db 1094 aagaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 507
 QY 508 gttcgggaatcatggtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 567
 Db 1154 aatcgggaatcatggtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 567
 QY 568 gctggggtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 627
 Db 1214 ccggggtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 627
 QY 628 gaaatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 667
 Db 1274 aagaatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 667
 QY 688 gacttgctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 747
 Db 1334 ggccttgctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 747
 QY 748 taaccggtatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 806
 Db 1394 taaccggtatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 806

QY 807 gtaacgaacttaacgagtaagtagaccgcttgaggactatgctcgcgaagaagtgaactc 866
 Db 1454 ctccggagactaaccggttaagtagaccgcttgaggactatgctcgcgaagaagtgaactc 866
 QY 867 aagaatgtagcgggtggtccgcgaagaagcggtagagcagcagcagcagcagcagcagcagcagc 926
 Db 1514 aagaatgtagcgggtggtccgcgaagaagcggtagagcagcagcagcagcagcagcagcagcagc 926
 QY 927 caaaacttaacttggtggtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 985
 Db 1574 gaagaacttaacttggtggtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 985
 QY 986 ggcagatt--cacaggtgctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1043
 Db 1634 ggaagcgtgtagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1043
 QY 1044 agtccgcgaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1100
 Db 1694 agtccgcgaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1100
 QY 1101 acgaactgctggtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1160
 Db 1754 ggaagcgtgctggtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1160
 QY 1161 tatgcaagggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1220
 Db 1814 taagagtaggctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1220
 QY 1221 ggaagc 1280
 Db 1874 gcaagc 1280
 QY 1281 aagtcgaatcagc 1340
 Db 1934 aagtcgaatcagc 1340
 QY 1341 gtacacacgcccgtac 1400
 Db 1994 gtacacacgcccgtac 1400
 QY 1401 taagagc 1459
 Db 2053 tcgggagggcgcttaaccacttgatgacgtggtgagcagcagcagcagcagcagcagcagcagcagcagc 1459
 RESULT 11
 ; US-09-897-516-3857
 ; Sequence 3857, Application US/09897516
 ; GENERAL INFORMATION:
 ; APPLICANT: Corbin, David R.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Huesing, Joseph E.
 ; APPLICANT: Krasomil-Osterfeld, Karina C.
 ; APPLICANT: Malvar, Thomas M.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Spiridonov, Sergei
 ; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
 ; FILE REFERENCE: 38-21(51847)B
 ; CURRENT APPLICATION NUMBER: US/09/897,516
 ; CURRENT FILING DATE: 2001-06-29
 ; PRIOR APPLICATION NUMBER: US 60/215, 161
 ; NUMBER OF SEQ ID NOS: 8409
 ; SEQ ID NO 3857
 ; LENGTH: 83080
 ; TYPE: DNA
 ; ORGANISM: Xenorhabdus sp.
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (10097)..(11929)
 ; US-09-897-516-3857

Query Match 51.7%: Score 763.4; DB 6: Length 83080;
Best Local Similarity 73.4%: Pred. No. 1.9e-252;
Matches 1100; Conservative 0; Mismatches 356; Indels 43; Gaps 8;

QY 1 gatcatggtcgaactaacgcttgccgagcgctcttaaacatgcaagtcgagcggttag 60
Db 616 gatcatggtcgaactaacgcttgccgagcgctcttaaacatgcaagtcgagcggttag 675

QY 61 caatac-----ctagcgagcaagcggttagtaacacgtgttaactct 102
Db 676 ggaagagctgtctgtcttgcttgacgagcggttagtaacatgtcttggttagctg 735

QY 103 cctccgagctcgttgataactcttcgaaagaaagttaacgataagctcctgttggctc 162
Db 736 ccgatctgaggggaaataacacacttggaacggttgtaatacagctgaacccttaagg-- 793

QY 163 acaagatttgaatgaatgaatattatgtcttgagatgagcccgagccgagctgaagtagt 222
Db 794 gcaaaagtgggagaccttcgagccctcaagccatcgatgaaccagatgggttagctagt 853

QY 223 tggtaggttaatgctcaacgaagcgagatcgtagccggtcgtgagaggtgtccggcc 282
Db 854 aggtggtgttaacggtctcaactcctagagcgagatccctagctgtgtctgagaggaatgaaccagcc 913

QY 283 acaatggaactgagacaagcgctcaactcctacggaagcgagccagcttaagatctgtctc 342
Db 914 acacttggagactggaacaacgagccagactccctacggaagcgagccagcttggtgataatgtgac 973

QY 343 aatgggggaaacccctgaaagcagacgacgacgtgaaacgaaagagcttcctgatatgtaa 402
Db 974 aatggggcgaagccttgatgacgagcagctgtgtatgaaagaagccttcggtgttgtaa 1033

QY 403 gtccaatgagcagagaa--?-----aataagcgaactgtgatagtgtact 447
Db 1034 gtacttccaggggagagagagcgctgaaagttaacgtcttccagatgtgactaacggc 1093

QY 448 gccataagcagcgctaaactacgtgacgagcgcggttaatactagtggtgcaagcgct 507
Db 1094 agaaagaagcagcgctaaactacgtgacgagcgcggttaatactagtggtgcaagcgct 1153

QY 508 gtcggaatcatgagcgctaaaggtgctgtagcggaatttgtaacgagctggtgaataact 567
Db 1154 aatcggaattactggtgctgaaagcgagcgagcggttcaattagtgatgtgaataacc 1213

QY 568 gcggtgctcaacccgtgctgctgacttgataactacaagctgtgagtttggtgagaggaagct 627
Db 1214 ccgggtctaaacctggtgaactgcatctaaagactgtgtgactagagcttcgttagagggggt 1273

QY 628 ggaattccaggtgtagcggtgaatgctgatatcttgaaagaaacacagctggcggaagcg 687
Db 1274 agaatctcaacgtgtagcggtgaatgctgatatcttgaaagaaacacagctggcggaagcg 1333

QY 688 gaattgctgctcaaaaactgacgcttgagcgacgaagcggtggttagtaaacggttaga 747
Db 1334 ggcgccctgagcgagagactgagctcaggttgcaaaagcggtgggagcaaacagattaga 1393

QY 748 taccggtgaatccagcgccctaaagcgtgtgtac--cagttgttggtggttttaaccctca 806
Db 1394 taccctgtgtgtccacggttaaacgattgtgacttggaaggtgtgtgcttgaaacggtggtg 1453

QY 807 gtaacgaacctcaacgagattagtagagcgcttggtggtgaactatgcttcgcaaggtgaactc 866
Db 1454 ctccgagagctcaacgaggttaagtagcagcgcttggtggtgaactatgcttcgcaaggtgaactc 1513

QY 867 aaaggaattgacgggggttcgcacaaagcggttgagcagtgtgttaattcaatgataatccc 926
Db 1514 aaaggaattgacgggggttcgcacaaagcggttgagcagtgtgttaattcaatgataatccc 1573

QY 927 caaaaactcaccctgggttgacatgtagatcgaatcat--gtagagataatagaaccttcg 985
Db 1574 gaagaaccttaaccttgaatcgaatcgaatccctcagaagatgagaggtggtccttcg 1633

QY 986 ggcagact--cacaggtgctcatggtgtgtcctcagctcgtctgtgagatgttggtta 1043
Db 1634 ggaagcgtagacagaggtgtgtcgaatggtcgtctcgaagctcgtgtgtgaaagtgttggtta 1633

QY 1044 agtcccgcaacgagcgcaacccctatactgtagtctgc---taacctaaagtgtgacactgt 1100
Db 1694 agtcccgcaacgagcgcaacccctatacttctgtgtgcagacagctgaatgtgtggaactcaa 1753

QY 1101 acgaacttgcggttgacaacaccggaaggaagcggtgagatgacgtgcaaatccatgacct 1160
Db 1754 gggagactgcggttgataaacccggaaggaaggttggtgagacgaatgacatcagctccct 1813

QY 1161 tatgtccagggcgcaacacggtgctacaaatggcgatataagaggttcgcgaactcgcaaga 1220
Db 1814 tacgagttaggtctacacacgtgtctatacaatggcagatataaagaagaagacactcgcgaga 1873

QY 1221 gggagctaatctctaaagaagctgcctccagctcggatgtgggtctctgcaactcgaccatg 1280
Db 1874 gcaagcggaactcataaagatctgtctgtagtcggatgtgagctgcaactcgactcactg 1933

QY 1281 aagtcgaatcgtctagtagatcgcggtatcagcatcgcggtgtgaataagttcccgacct 1340
Db 1934 aagtcgaatcgtctagtagatcgcggtatcagcatcgcggtgtgaataagttcccgacct 1993

QY 1341 gtacacaccgcccgtcacacaccctggaatggggagacccgaaagtgtgtctgtgtacgg 1400
Db 1994 gtacacaccgcccgtcacacacatggaatgggtgtgtgcaaaaagaatag--gtagcttaact 2052

QY 1401 taaggaagacagactaactaaggctgaaaactcgtaaaggggtgaatctgtaaaaggtacc 1459
Db 2053 tcgggagggcgcttaacacttltgattcatgacttggtgaggtgaagtcgtaaagaaggtaac 2111

RESULT 12
US-09-897-516-3858
: Sequence 3858, Application US/09897516
: GENERAL INFORMATION:
: APPLICANT: Corbin, David R.
: APPLICANT: Goldman, Barry S.
: APPLICANT: Hinkle, Gregory J.
: APPLICANT: Huesling, Joseph E.
: APPLICANT: Krasomil-Osterfeld, Karina C.
: APPLICANT: Malvar, Thomas M.
: APPLICANT: Slater, Steven C.
: APPLICANT: Spiridonov, Sergei
: TITLE OR INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
: FILE REFERENCE: 38-21(51847)B
: CURRENT APPLICATION NUMBER: US/09/897,516
: PRIOR FILING DATE: 2001-06-29
: PRIOR APPLICATION NUMBER: US 60/215, 161
: NUMBER OF SEQ ID NOS: 8409
: SEQ ID NO 3858
: LENGTH: 83080
: TYPE: DNA
: ORGANISM: Xenorhabdus sp.
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (17551)..(19197)
US-09-897-516-3858

Query Match 51.7%: Score 763.4; DB 6: Length 83080;
Best Local Similarity 73.4%: Pred. No. 1.9e-252;
Matches 1100; Conservative 0; Mismatches 356; Indels 43; Gaps 8;

QY 1 gatcatggtcgaactaacgcttgccgagcgctcttaaacatgcaagtcgagcggttag 60
Db 616 gatcatggtcgaactaacgcttgccgagcgctcttaaacatgcaagtcgagcggttag 675

QY 61 caatac-----ctagcgagcaagcggttagtaacacgtgttaactct 102

```
Dh 676 ggaagcagctctgctctgtctgtacgagcgcgagcggtgaaatgctgtgggacgtc 735
Qy 103 cctccgagctctggataactctccgaaagaaagctaatccgagatgcctgttgacc 162
Db 736 ccgagctgggggagataaacacttgaaacggtgtgctaatccgcatgaccttaagaa-- 793
Qy 163 acaagaattgataagtaaaagattatctgtgtgagaatgagccgagccgacctatgagc 222
Db 794 gcaaaagtgggggagaccttcgggctcagccatcgatcgatgaacccaagtggatgagta 853
Qy 223 tggtagtgtaatgtgtccacaaagcgagatcgtagccggtcctgagagaggtgtccggcc 282
Db 854 aagtgaggtaaacggtccactctagtcgacatcccttagctgtgtctgagagatgacacagc 913
Qy 283 acaatggaactgagacacggtccatctcctctagcggagcagcagatgaagtaactgtgtc 342
Db 914 acaatggagcttgagacacggtccacagctcctctacgagagcgagcaggtggggaatatgcac 973
Qy 343 aatgggggaaacccctgaagcagcgagcggtgtgaacgaagaaggtctctcgatgtgaa 402
Db 974 aatgggagcaagcctgtatgacgcatgcgcggtgtatgaagaagccttcgggtgtgaa 1033
Qy 403 gtctatgagcagga-----aataagacgcatgtgatgtgacacct 447
Db 1034 gtaacttcagcggggaggaagcggtgaagtgtgaatacgttccgactgacgttaccgcgc 1093
Qy 448 gcttaagcaccggttaactacgtgtgcacagcgcggtgaatacgtatgtgtcaagcgtt 507
Db 1094 agaaagagcagcggttaactcgtgtccagcagcggtgtgaatacgaaggtgtgaagcgtt 1153
Qy 508 gtctcggaatcattgtggcgtaaaaggtgtgcgttagcgagattgttaagtcaaggtgtgaaact 567
Db 1154 aatcggaattactggcggttaaaagcgacgacggtgtgcaatlaagttagatgtgaatcc 1213
Qy 568 ggagggtcaacccggtgtgccttgcaacttgaaactaaactgtgtgaggtttgggaagggaagt 627
Db 1214 ccggggttaactcgtggaactgtcaactaaagcgtgtgactaagagctcccgtaagggggggt 1273
Qy 628 ggaatccaggtgtcagcggttgaatgtcgtagatcctcgtgagaaacccagctgtgcgagagc 687
Db 1274 agaatccacggtgttagcgtgtgaatgtcgttagagatgttgtaggaatacgcgtgtgcgagagc 1333
Qy 688 gactgtgtgtcccaaacctgtagcgtgtgagcagcaagcggtgtggttagtaaacgggattaga 747
Db 1334 ggcacctgtgaagaaagactgtagcgtgtcgaagcggtgggagcaaacaggtattaga 1393
Qy 748 taaccggtgaatccagcggtccttaacgtgtgttac--cagttgttgggggttttaaccctca 806
Db 1394 taccctgtgtagtcacacggtcgttaaacgattgtcgacttggaaggttgtgtgcttgagccggtg 1453
Qy 807 gtaacgaaccttaagattaaatagacgcgtgtgggagatactcgcgaagatgaaactc 866
Db 1454 ctctcgagactaagcgtttaaagtcgacgcgtcgtggagtagcgttcgaagattaaactc 1513
Qy 867 aagaagaattgaaagggtgtccgacaaagcggtgtgagatgtgtgttaattcgtatgatacc 926
Db 1514 aaaaatgaaatggaggggtcccgacaaagcggtgtgagcattgtgttaattcgtatgacgcgc 1573
Qy 927 caaaaactcactgtggtctgtgacatgacatcgtatcat--gttagaaatataatgaccttcg 985
Db 1574 gaaagaaccttaactactctctgtgacatccagaaactcctcaagagatgaaggatgtaccttcg 1633
Qy 986 ggcagatt--caacaggtgtgtgacatgtgtcgtcagcgtgtgtcgtgtgagatgttggtta 1043
Db 1634 gaaacgcttgagacaggtgtgtcgtacatgtgtcgtcagctcgtgtgtgtgaaatgttggtta 1693
Qy 1044 agtccgagacagcgcaaacctctacgtatgttgc---taacttaagtgtggcaactgtgt 1100
Db 1694 agtccgacagcagcgcaaacctctacgttctgttcagacagctaaatgtgttggaactcaaa 1753
Qy 1101 acaaaactcgcgtgtgacaaacggaggaagcggtgtgagatgacgtcaaatctcctatgacct 1160
Db 1754 gggagagctcgcgtgtgataaacggaggaaggtgtgggagatgacgtcaaatcattacgtacct 1813
```

```
Qy 1161 tatgtccagggccacagacgtgtctactatgtgcccgatatacagaaggtgtcccaactcgaaga 1220
Db 1814 tacgaatgaagggtacacacagcgttactaaatgtgcagatacaaaagaagcgacctcgcga 1873
Qy 1221 gggagctaatctctaaagtcgtccagcttcgagattgtgggtctcgaactcgaaccccatg 1280
Db 1874 gcaagcggaactataaagtcgtcgttagtcggtattgttgatgttcgaactcgaactcgtccatg 1933
Qy 1281 aagtggatcgtctagtaatcgcggtatcagcatgtccgcggtgtgaatacgttcccgacct 1340
Db 1934 aagtcggaatcgtctagtaatcgttagatcagaaatgcatcgtgtgatatcgttcccgacct 1993
Qy 1341 gtaacacgcccgttaccacccactgagtgaggagaccccggaaggtgtctgtgtgaacg 1400
Db 1994 gtaacacgcccgttaccacccatcgtgaggtgtggtgtgcaaaagaagtag--gtagcttaact 2052
Qy 1401 taaggagacagactactaaagtgtaaactcgttaaaagggtgtgaagtcgtatacaaggtacc 1459
Db 2053 tcggagagcggtcttaccacttgtgtatcgtatgactgtgggtgaagtcgtatacaaggtacc 2111
```

```
RESULT 13
US-09-897-516-3859
; Sequence 3859, Application US/09897516
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Huesing, Joseph E.
; APPLICANT: Krasomil-Osterfeld, Karlna C.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Slater, Steven C.
; APPLICANT: Spiridonov, Sergei
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51847)B
; CURRENT APPLICATION NUMBER: US/09/897,516
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215, 161
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 8409
; SEQ ID NO 3859
; LENGTH: 83080
; TYPE: DNA
; ORGANISM: Xenorhabdus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (37881)..(38108)
; US-09-897-516-3859
```

```
Query Match 51.7%; Score 763.4; DB 6; Length 83080;
Best Local Similarity 73.4%; Pred. No. 1.9e-252;
Matches 1100; Conservative 0; Mismatches 356; Indels 43; Gaps 8;

Qy 1 gatcatgtctcgaactaaacgctgtgcgcgctcttaacaatgcaagtcagcggtgtag 60
Db 616 gatcatgtctcagattgtgaacgctgtgcgcgagcggtcaacatgcaagtcagcggtgaaca 675
Qy 61 catatc-----ctagcggcggaaggggtgtgtaacaacgtgtgtaactt 102
Db 676 ggaagaagcttgtctgtctgtcgtgaacgagcggtggaaggtgtgtaacttctgtggaatctg 735
Qy 103 cctccgagctctggataactctccgaaagaaagctcaataccgagatgcctgttgatc 162
Db 736 ccgagatgggggagataaacacactcggaacggtgtgcttaataccgcatgaccttaagaa-- 793
Qy 163 acaagaattgataagtaaaagattatctgtgtgagaatgagccgagccggtatgagtagt 222
Db 794 gcaaaagtggggagaccttcgggctcagccatcgatcgatgaacccaagatgtgagtagt 853
Qy 223 tggtagtgtaatgtgtccacaaagcggtgtgagatcgtgtgagaggtgtgtccggcc 282
Db 2053 tggtagtgtaatgtgtccacaaagcggtgtgagatcgtgtgagaggtgtgtccggcc 282
```



```
Db 854 agtgggaaacggtcaccatgagcgacgataccatgctggtctgtgagagatgacagcc 913
Oy 283 acaatggagacgtgagacagcgctcattctctagaggagcgagcaatlaagatctgtc 342
Db 914 acaatggagacgtgagacagcgctcattctctagaggagcgagcaatlaagatctgtc 973
Oy 343 aatggggaacacccctgaagcagcgacgacgctgtgaacgaagaagctctcgtatgttaa 402
Db 974 aatggggaacacccctgaagcagcgacgacgctgtgtatgaagaagcctcgggtgttaa 1033
Oy 403 gtctaatggcagagaa-----aataagcagcaatgtatgttacct 447
Db 1034 gtactttcagcgaggaggaagcggtgaagtgaattacatcttccagatgtgacccgc 1093
Oy 448 gccataagcaccggttaactcgttgcacagcgccggtlaactatggtggaagcggt 507
Db 1094 agaaagaacacccggttaactcgttgcacagcgccggtlaactatggtggaagcggt 1153
Oy 508 gtctggaatcattggcgctaaagggtgctgtagcgagattgtlaagtcaggtgtgaact 567
Db 1154 aatcggaattactggtggttaagcgacgacgacgacggtcaatlaagtgatgtgaactc 1213
Oy 568 gcgggtcacaacccgtggtcgtcactgtgaactacaagtcgtgagtttgaggagcgaggt 627
Db 1214 ccgggtcacaacccgtggtcgtcactgtgaactacaagtcgtgagtttgaggagcgaggt 1273
Oy 628 ggaattcaggtgtgagcggtgaatgctgtagatctctgaggagaacacgagtgagagcg 687
Db 1274 agaatccacgctgtagcggtgaaatgctgtagatctctgaggagaacacgagtgagagcg 1333
Oy 688 gactgtcgtgctcaaaatctgacgtctgagcaagacgctggtgtagtaacgaggtatga 747
Db 1334 ggcgccctcgagcaagacgtcagtcaggtgcaagacgctggtgtagtaacgaggtatga 1393
Oy 748 taaccggttaactcaacgctcctaacaagtcgttac--cagttgttgagggttttaacctca 806
Db 1394 taaccggttaactcaacgctcctaacaagtcgttac--cagttgttgagggttttaacctca 1453
Oy 807 gtaacgaacctaacggttaagtagacggtcggtgaggtatgctgcaagaagtgaacct 866
Db 1454 ctcccgaggtcctaacggttaagtagacggtcggtgaggtatgctgcaagaagtgaacct 1513
Oy 867 aaaggaattcgacgggtgctcgcacaaacgctgtagagcatgtgttaattcgaatgaacc 926
Db 1514 aaatgaattcgacgggtgctcgcacaaacgctgtagagcatgtgttttaattcgaatgaacc 1573
Oy 927 caaaaactcaactgtggttgaatgacatgacatcat--gtcagaatatatgaacctcg 985
Db 1574 gaaagaactcaactgtggttgaatgacatgacatcat--gtcagaatatatgaacctcg 1633
Oy 986 ggaagatc--cacaggtgctgcatgtgtgtcagctcgtgtcgtgagatgttggtta 1043
Db 1634 ggaagatc--cacaggtgctgcatgtgtgtcagctcgtgtcgtgagatgttggtta 1693
Oy 1044 agtcccgcaacgagcgcaaacctcctatgtgtgctgctgctgctgctgctgctgctgct 1100
Db 1694 agtcccgcaacgagcgcaaacctcctatgtgtgctgctgctgctgctgctgctgctgct 1753
Oy 1101 acgaacatcgccggtgtagcaaacacggagagcggtgtagatgacatcctcattgtgctt 1160
Db 1754 gggagatcgccggtgtagcaaacacggagagcggtgtagatgacatcctcattgtgctt 1813
Oy 1161 tatgccaagcgccacacgctgtctacaatgtagcgaatacgaaggtgcgcaactcgaaga 1220
Db 1814 taaggtatgtaggtgtacacgctgtgtcacaatgtagcgaatacgaaggtgcgcaactcgaaga 1873
Oy 1221 gggagatcattctaaagtcgtgtccagttcgtgattgtgtgtcgtcaactcgaacctga 1280
Db 1874 gcaagcggaactcattaaagtcgtgtccagttcgtgattgtgtgtcgtcaactcgaacctga 1933
Oy 1281 aagtcggaatcgctgtagatcgtagcgaatcagcgtgagcgtgtaataagttcccgagact 1340
Db 1934 aagtcggaatcgctgtagatcgtagcgaatcagcgtgagcgtgtaataagttcccgagact 1993
```

```
Oy 1341 gtacacaccgcccgtacacaccacccctgaatggggagacccggaagtgtctgttaacg 1400
Db 1994 gtacacaccgcccgtacacaccacccctgaatggggagacccggaagtgtctgttaac 2052
Oy 1401 taaggaacagactaactaagtgtaaacctcgtlaaagggtgtaagtcgttaacaagttacc 1459
Db 2053 tcggagagcgcttaccactctgtatcattcagctggtgggtgtaagtcgttaacaagttacc 2111

RESULT 14
US-09-897-516-3860
: Sequence 3860, Application US/09897516
: GENERAL INFORMATION:
: APPLICANT: Corbin, David R.
: APPLICANT: Goldman, Barry S.
: APPLICANT: Hinkle, Gregory J.
: APPLICANT: Huesling, Joseph E.
: APPLICANT: Krasomil-Osterfeld, Karina C.
: APPLICANT: Malvar, Thomas M.
: APPLICANT: Slater, Steven C.
: APPLICANT: Spiridonov, Sergei
: TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
: FILE REFERENCE: 38-21(51847)B
: CURRENT APPLICATION NUMBER: US/09/897,516
: CURRENT FILING DATE: 2001-06-29
: PRIOR APPLICATION NUMBER: US 60/215, 161
: PRIOR FILING DATE: 2000-06-30
: NUMBER OF SEQ ID NOS: 8409
: SEQ ID NO 3860
: LENGTH: 83080
: TYPE: DNA
: ORGANISM: Xenorhabdus sp.
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (40567)..(40932)
US-09-897-516-3860

Query Match 51.7% Score 763.4; DB 6; Length 83080;
Best Local Similarity 73.4%; Pred. No. 1.9e-252;
Matches 1100; Conservative 0; Mismatches 356; Indels 43; Gaps 8;

Oy 1 gatcatgctcagaactaagctggtcggtcgtcttaacaatgcaatgcagcggtgtag 60
Db 616 gatcatgctcagaactaagctggtcggtcgtcttaacaatgcaatgcagcggtgtag 675
Oy 61 caatac-----ctagcgacgacgggtgtagtaacaagtgttaactt 102
Db 676 ggaagcagcttgctgtcttctgtcagagcggtgtagtaagtgtcgtggtgtagctg 735
Oy 103 cctcagatctggagtaactcttcggaagaagtaacgggtatcctcgttggtgtagc 162
Db 736 ccgagtggaagggtgaataacacgtgtaaacgcatgacacgcatgacacctaaga-- 793
Oy 163 acaagatttgatagtgtaagaattatctgtctgtagatagcccggtcgagattagtagt 222
Db 794 gcaaatgtggggaactcttggtcgtcagcagcagatgaaccagatgtggttagtagt 853
Oy 223 tggtaggtatagtgctcaacaaagcgacgacgtgtagcgtcggtgtagaggtgttcggcc 282
Db 854 aggtgggttaagtgctcaacaaagcgacgacgtgtagcgtcggtgtagaggtgttcggcc 913
Oy 283 acaatggagacgtgagacagcgctcattctctagaggagcgagcaatlaagatctgtc 342
Db 914 acaatggagacgtgagacagcgctcattctctagaggagcgagcaatlaagatctgtc 973
Oy 343 aatggggaacacccctgaagcagcgacgacgctgtgaacgaagaagctctcgtatgttaa 402
Db 974 aatggggaacacccctgaagcagcgacgacgctgtgtatgaagaagcctcgggtgttaa 1033
Oy 403 gtctaatggcagagaa-----aataagcagcaatgtatgttacct 447
```

Dh	1034	gtatcttcagtcggtgggaggaagcgtgaaattgtaataacgtttcaacgatctgaagcttaaccgc	1093
Oy	448	gctctaaagcacccggtcctaactacgtctccagcacgcgcgtgtaalaacgtatgctgaacgctt	507
Dh	1094	agaagaagcaccccgctctaactccgtctcccaacacacgcgcgtgtaatactcgaaggtgtcaacgctt	1153
Oy	508	gttcgcgaatactatgggcgttaaaggtgtgcgttagcggatttgtaactcaggtgtgtaaaact	567
Dh	1154	aattcggaaattcaactcggcgttaaaagcgaacgcgaagcgggttcataaattgaatgatacgtgtaaatcc	1213
Oy	568	gcgggcgtcaaaccccgctgcgcctcgtcaactgtaaaacaaagtcctgtaaggtcttgtagaaggaacgt	627
Dh	1214	ccgggctctaaccctcgggaactcgtacataagaactcgtttgaactagatgctctagtaggggggt	1273
Oy	628	ggaattccaggtgttagcgtgtgaaaatgcgtatagatctctgtaggaagacaacagatgtgcgaagc	687
Dh	1274	agaattccaaagtttagcgtgtgaaatctcgtataagatgtgttgtaggaataataccgtgtgcgaagc	1333
Oy	688	gactctgtgcctcaaaaactgtacgcctgtggcacaagaacgcttgtagtaaacgggattaga	747
Dh	1334	ggccccctcggacgaagaactgtgaacgtctcaaggtgcgaagaacgctgtgggggcacaacgaagattaga	1393
Oy	748	taccctgttatccacacgcgccttaaacgtctgtctac--cagttcttggtgggtctttaaccctca	806
Dh	1394	taccctgttatccacacgcgtttaaagatgtcgaactgtggaaggtgtgtgcctcttgagccgtgg	1453
Oy	807	gttaacggaacttaacggaattgaattgtagccgcgcgtgggagactatgctgcgaaggtgtgaactc	866
Dh	1454	cttcctcggagacttaaacgcgtttaaagtctgaacccgcctgtggggagatgaacgcttcgaacgaattaaactc	1513
Oy	867	aaaggaattcgaacgggggtctccgcacaagcgcgtgtgagacatgttgatttaactcgaatgataacc	926
Dh	1514	aaatgaaattgagcgggtggccgcacaagcgcgtgtgaggaatgtgttttaattcgaatgaacgc	1573
Oy	927	caaaaacctcaactcgtggcttggaatgatacgtgaatcat--gtagaatatatagagctctcg	985
Dh	1574	gaagaacacttaacttaactctctgtgaatcccaagcgaatcccttaagaagatgtagggagtgctctcg	1633
Oy	986	ggcagcat--cacagtggtcgtgcatgtgtgtgcgtcaacatcgtctctctgtagatgtgtgtta	1043
Dh	1634	ggaaacgtctgagacaagcgtgtgcgtcaatgtcgtgtcgtctgcagctgtgtgtgtaaaagtgtgtgtta	1693
Oy	1044	agtcctgcgaacgagcgaacccctcatcgtatgttc--taacctaaagtgtggcaactggt	1100
Dh	1694	agtcctgcgaacgagcgaacccctcatccttctgtgtgcgaacgaatgatacgtgtgaaactcaa	1753
Oy	1101	acggaactgcgcgtgtgacaacccggtgaggaagcgggtgatacgtccaatccatcgaatgcctt	1160
Dh	1754	ggggagactgcgcgtgtgataaacccggaaggaaggtgtgggatacgtccaagtatcatatgcctc	1813
Oy	1161	tatctccagggcccaacacatgtctatacatgtccgatatacgaaggggtctgcgaactgcgaaga	1220
Dh	1814	tacgaagttagggctatacacatcgtctatacatgtgcgaatatacaagaagaagagactcgcgaga	1873
Oy	1221	ggggagctaatctctcaaaagtctcgttcccagttcggattcgggtctgcgaactcgaaccctatg	1280
Dh	1874	gcaagcggaaatctcataaagctctgtctgtatccggattctggaattcgtcaactcgaactcctatc	1933
Oy	1281	aagtcggaatctgcgtatgaataatccgcgtatccaacatgcgcgcgtgtgaatacgttcccgaactt	1340
Dh	1934	aagtcggaatctgcgtatgaataatcgtatgaatacgaatagcgaatgatacgttcccgggcctt	1993
Oy	1341	gtacaacacgcgcgtctcaacacactatgcttggtggagacaccgcgaagatgtcttctgttaacg	1400
Dh	1994	gtacaacacgcgcgtctcaacacatgtggaggtgtgtgtcgaanaagaagtatg--gtagcttaacct	2055
Oy	1401	taaggtgagacagactctaaggtgtgaactcgtttaaggggtgtgaagctgtatacaaggttacc	1459
Dh	2053	tcgggtgaggtcgtctacaacttctgtatctacatgtcgtgggtgtgaagctgtatacaaggttacc	2111

```

? Sequence 3861, Application US/09897516
? GENERAL INFORMATION:
? APPLICANT: Corbin, David R.
? APPLICANT: Goldman, Barry S.
? APPLICANT: Hinkle, Gregory J.
? APPLICANT: Huesing, Joseph E.
? APPLICANT: Krasomil-Osterfeld, Karina C.
? APPLICANT: Malvar, Thomas M.
? APPLICANT: Slater, Steven C.
? APPLICANT: Spiridonov, Sergei
? TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
? FILE REFERENCE: 38-21(51847)B
? CURRENT APPLICATION NUMBER: US/09/897,516
? CURRENT FILING DATE: 2001-06-29
? PRIOR APPLICATION NUMBER: US 60/215, 161
? PRIOR FILING DATE: 2000-06-30
? NUMBER OF SEQ. ID NOS.: 8409
? SEQ. ID NO 3861
? LENGTH: 83080
? TYPE: DNA
? ORGANISM: Xenorhabdus sp.
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (41695)..(42120)
? US-09-897-516-3861

```

Query Match	51.7%;	Score 763.4;	DB 6;	Length 83080;
Best Local Similarity	73.4%;	Pred. No. 1.9e-252;		
Matches 1100; Conservative	0;	Mismatches 356;	Indels 43;	Gaps 8

Oy	1	gatactggtccaaataactgacgtctggccggcgcgctcttaaacatgtaactgagcggttag	60
Db	616	gattcaatggctccaaatattgaaacgcgtgcggcgaagccttaacacatgcagtcagcggtacaa	67
Oy	61	caatac-----ctagcggcgaaacgggttgagttaacacgttgtaactctt	10
Db	676	ggaagcagcgtctgcgtcttgcttgagaaagacgvcgagcvggttgagttaagtctggcgatctg	73
Oy	103	ctctccggtcttgtagtaactcttcggaaaggaaagcttaattacggatactgcgttgtagtc	16
Db	736	cccgatgagggaggtataacacacttgtaaaagagctggtcattacatcgaaatgactctcaagga	79
Oy	163	acaaagatttatatggtatgaagatttatctgtcttgagatgagaccgcggccgattatgctagt	22
Db	794	gcaaaagggtggagactcttcggcgtccatcgatcagatgaaacccagatgagattatgactagt	85
Oy	223	tgtgtgagatgatactgctccaaagagcgagatctgtatgacgcgctcgaaagaggttgtctgcgc	28
Db	854	aggttgggtgtaaacgcgtccacactatgagcgaatctccaaagctctgcgtgagatgaaatgaacagcc	91
Oy	283	acaaatgaaacttgtagacaacgctccatactctctacgaggagcagatgtgaaagatactctgtctc	34
Db	914	acactgtagacttgagacaacgcccagaatctctacggtgaggcagcagttgggaatatgtcac	97
Oy	343	aatggggggaaccctgtgaagcagcagcgcgcgctgtgaagaaagaaaggtctctcggaattgtaa	40
Db	974	aatgggcgcgaagccttgatgcagacatctgcgcgcgtgtgataaaagagcctctcggttgttaa	10
Oy	403	gttactatgagcaggaata-----aataaacagcaatgtgcatgctgtact	44
Db	1034	gtactcttcagcgggggagagcggtgaaagtgtgaatacgttttcaagatgtgcagcttaccgcg	10
Oy	448	gctctaaagcacccgcttaactacgtctgcacagcagcgggttaatacgtatggttcgaagcgtt	50
Db	1094	agaagaagcacccgctataactcgcgttccacagcagcccggttaatacggaggggttgcacaagcgtt	11
Oy	508	gtctcggaatattgggcgttaaaaggtggtgcgtgaagcggattgtgaatcaggltgcaaacct	56
Db	1154	aatcgggaattactcggggtgttaaaagcgcacagcagcggttcaattaaagttaagatgtgaaatcc	12
Oy	568	gcgggcctcaaacccgttgacctgacattgaaactacaagtcttgagatttgagagagcaagt	62

Db 1214 ccgggcttaacctggaactgcatacctaagactggttgactagactcctcgttagagggggt 1273
Qy 628 ggaattccaggtgtgagcgttgaatctgtaatactcgtgagaacaccagtcggcgaagc 687
Db 1274 agaatlcacacgtttagcgttgaatctgtaagagatgttgaggaaataccgttgcgaagc 1333
Qy 688 gacttgcgtgctcaaaactgacgttgaagcgaagaagcgttggtagtaataccggataga 747
Db 1334 gggccctcggcgaagaactgcgtcagtcgtaagcgttggagcgttggagcaaacagataga 1393
Qy 748 taaccggtaatacgaagcctaagcgttgcac-cagttgttgggggtttaacctca 806
Db 1394 taacctgttagtccacgcgttaaacgacttgacttgagggttcttgcccttgagccgtg 1453
Qy 807 gtaacgaacctaaagatlaagtagaccgcttgggactatgcttcgcaaggtgaactc 866
Db 1454 ctccggagctaaacgctlaagctgacccgcttgggagtagcgttcgcaagatlaaaactc 1513
Qy 867 aaaggaattgacgggggtcgcacaaagcgtggaagcgttggatlaactcgaatacacc 926
Db 1514 aaatgaattgacgggggtcgcacaaagcgttggagcgttggatlaactcgaatacacc 1573
Qy 927 caaaaacctcaactggtgcttgacatgactgaatcat-gtagagataatgagccttcg 985
Db 1574 gaagaaaccttaactcttgaacatccagcgaatccttcagagatgaaggagtgccctcg 1633
Qy 986 ggcagatt--cacagggtgctgacgttgcgtacagctcgttcgttgagatgttgggtta 1043
Db 1634 ggaacgctgagacaggtgtcgtacgtgctgcgtcagcgtcgttctgtaaatgttgggtta 1693
Qy 1044 agtcccgcaagcgaagcgaacctatacgtatcgttcg---taacctaaattggcactgt 1100
Db 1694 agtcccgcaagcgaagcgaacctataccttctgttcgcagcagcgtatgttggaaactcaa 1753
Qy 1101 acgaactgcccgttgcacaaaccggaaggaaggcgggagtgacgtlcaaatcctcatgacct 1160
Db 1754 gggagactgcccgttgcataaacccggaaggaaggcgggagtgacgtlcaaatcctcatgacct 1813
Qy 1161 tatgtccaggggcacacacgttgcataaaatggccgatacagaagggtcggcaactcgcgaaga 1220
Db 1814 taagagtagggcttaacacacgttgcataaaatggccgatacagaagggtcggcaactcgcgaaga 1873
Qy 1221 gggagctaatctctaaagctcgttccagctcggatltggggcttgcgaactcgaacctca 1280
Db 1874 gcaagcgaactcatalaaagctcgttccgtagtccgagltggagcttgcgaactcgaacctca 1933
Qy 1281 aagtcggatcgtcgtatgtagtgcgagatcagcatgcgcgttgaatacgttcccggaact 1340
Db 1934 aagtcggatcgtcgtatgtagtgcgagatcagcatgcgcgttgaatacgttcccggaact 1993
Qy 1341 gtacacacccgcttcaacacacacccgtgagtgaggacccggaagtgtccttgttaaccg 1400
Db 1994 gtacacacccgcttcaacacacacccgtgagtgaggatgttgcgaagaagatag-gtagcttaacct 2052
Qy 1401 taaggagacagactactaaggttgaactcgttaaaagggttgaagtgcgttaacaaggtacc 1459
Db 2053 tcgggagggcgttaccacttctgtatcattgacgtgggttgaggtcgttaacaaggttaac 2111

Search completed: December 15, 2001, 05:58:16
Job time: 12259 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 05:42:50 : Search time 7962.77 Seconds
(without alignments)
48.423 Million cell updates/sec

Title: US-09-380-826a-2

Perfect score: 22
Sequence: 1 tfttgatcaccaagattgata 22

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 17159718 seqs, 8763200856 residues

Total number of hits satisfying chosen parameters: 34319436

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_MA_Main:*

```
1: /cgn2_6/ptodata/2/pna/US07PCTUS.COMB.seq:*
2: /cgn2_6/ptodata/2/pna/US06.COMB.seq:*
3: /cgn2_6/ptodata/2/pna/US07.COMB.seq:*
4: /cgn2_6/ptodata/2/pna/US081.COMB.seq:*
5: /cgn2_6/ptodata/2/pna/US081.COMB.seq:*
6: /cgn2_6/ptodata/2/pna/US083.COMB.seq:*
7: /cgn2_6/ptodata/2/pna/US083.COMB.seq:*
8: /cgn2_6/ptodata/2/pna/US084.COMB.seq:*
9: /cgn2_6/ptodata/2/pna/US085.COMB.seq:*
10: /cgn2_6/ptodata/2/pna/US086.COMB.seq:*
11: /cgn2_6/ptodata/2/pna/US087.COMB.seq:*
12: /cgn2_6/ptodata/2/pna/US088.COMB.seq:*
13: /cgn2_6/ptodata/2/pna/US089.COMB.seq:*
14: /cgn2_6/ptodata/2/pna/US090.COMB.seq:*
15: /cgn2_6/ptodata/2/pna/US091.COMB.seq:*
16: /cgn2_6/ptodata/2/pna/US093.COMB.seq:*
17: /cgn2_6/ptodata/2/pna/US093.COMB.seq:*
18: /cgn2_6/ptodata/2/pna/US094.COMB.seq:*
19: /cgn2_6/ptodata/2/pna/US095.COMB.seq:*
20: /cgn2_6/ptodata/2/pna/US095.COMB.seq:*
21: /cgn2_6/ptodata/2/pna/US095.COMB.seq:*
22: /cgn2_6/ptodata/2/pna/US095.COMB.seq:*
23: /cgn2_6/ptodata/2/pna/US096.COMB.seq:*
24: /cgn2_6/ptodata/2/pna/US096.COMB.seq:*
25: /cgn2_6/ptodata/2/pna/US096.COMB.seq:*
26: /cgn2_6/ptodata/2/pna/US096.COMB.seq:*
27: /cgn2_6/ptodata/2/pna/US096.COMB.seq:*
28: /cgn2_6/ptodata/2/pna/US096.COMB.seq:*
29: /cgn2_6/ptodata/2/pna/US097.COMB.seq:*
30: /cgn2_6/ptodata/2/pna/US097.COMB.seq:*
31: /cgn2_6/ptodata/2/pna/US098.COMB.seq:*
32: /cgn2_6/ptodata/2/pna/US098.COMB.seq:*
33: /cgn2_6/ptodata/2/pna/US098.COMB.seq:*
34: /cgn2_6/ptodata/2/pna/US098.COMB.seq:*
35: /cgn2_6/ptodata/2/pna/US098.COMB.seq:*
36: /cgn2_6/ptodata/2/pna/US098.COMB.seq:*
37: /cgn2_6/ptodata/2/pna/US098.COMB.seq:*
38: /cgn2_6/ptodata/2/pna/US098.COMB.seq:*
39: /cgn2_6/ptodata/2/pna/US098.COMB.seq:*
40: /cgn2_6/ptodata/2/pna/US098.COMB.seq:*
41: /cgn2_6/ptodata/2/pna/US098.COMB.seq:*
42: /cgn2_6/ptodata/2/pna/US098.COMB.seq:*
43: /cgn2_6/ptodata/2/pna/US098.COMB.seq:*
```

SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	% Match	Query length	ID	Description
1	22	100.0	22	US-09-380-826a-2	Sequence 2, Appl
2	22	100.0	17	US-09-380-826a-7	Sequence 7, Appl
3	22	100.0	1477	US-09-380-826a-1	Sequence 1, Appl
4	18.8	85.5	420	US-09-380-826a-11944	Sequence 11944, A
5	18.8	85.5	530	US-09-649-165a-7425	Sequence 7425, Ap
6	18.8	85.5	581	US-09-654-016-175385	Sequence 175385,
7	18.8	85.5	581	US-09-654-016-175385	Sequence 175385,
8	18.8	85.5	649	US-09-565-309a-20908	Sequence 20908, A
9	18.8	85.5	1069	US-09-565-309a-57118	Sequence 57118, A
10	18.8	85.5	1072	US-09-565-309a-49885	Sequence 49885, A
11	18.8	85.5	1128	US-09-654-016-117747	Sequence 117747,
12	18.8	85.5	1429	US-09-654-016-11564	Sequence 11564, A
13	18.8	85.5	5874	US-09-404-520-5289	Sequence 5289, A
14	18.8	85.5	92624	US-09-534-859-352	Sequence 352, App
15	18.8	85.5	92624	US-09-803-736-352	Sequence 148, App
16	18.4	83.6	2101	US-09-143-993-148	Sequence 27704, A
17	18.4	83.6	2463	US-09-191-637-27704	Sequence 22313, A
18	18.4	83.6	2463	US-09-191-637-27704	Sequence 18072, A
19	18.4	83.6	2486	US-09-167-217-19072	Sequence 19008, A
20	18.4	83.6	3442	US-09-143-993-650	Sequence 650, App
21	18.4	83.6	3831	US-09-167-217-18350	Sequence 18350, A
22	18.4	83.6	4637	US-09-619-049-937	Sequence 937, App
23	18.4	83.6	4637	US-09-171-637-1565	Sequence 1465, App
24	18.4	83.6	4637	US-09-173-464-15063	Sequence 15063, A
25	18.4	83.6	4637	US-09-191-637-18351	Sequence 18351, A
26	18.4	83.6	4637	US-09-191-637-18351	Sequence 14523, A
27	18.4	83.6	4637	US-09-167-217-18359	Sequence 18359, A
28	18.4	83.6	53867	US-09-528-237a-1225	Sequence 1225, Ap
29	18.4	83.6	206	US-09-654-016-180482	Sequence 180482,
30	18.4	83.6	206	US-09-654-016-180482	Sequence 180482,
31	17.8	80.9	403	US-09-521-640-103292	Sequence 103292,
32	17.8	80.9	2084	PCT-US00-00724-1291	Sequence 1291, Ap
33	17.8	80.9	85992	US-09-534-859-802	Sequence 802, App
34	17.8	80.9	85992	US-09-803-736-802	Sequence 802, App
35	17.8	80.9	59261	US-09-534-859-23	Sequence 23, Appl
36	17.4	79.1	59261	US-09-803-736-23	Sequence 951, App
37	17.4	79.1	81662	US-09-534-859-951	Sequence 951, App
38	17.4	79.1	81662	US-09-803-736-951	Sequence 185, App
39	17.4	79.1	242	US-60-185-216-185	
40	17.2	78.2			
41					

42 17.2 78.2 249 24 US-09-638-258-520 Sequence 520, App
43 17.2 78.2 249 32 US-09-924-035-520 Sequence 520, App
44 17.2 78.2 278 25 US-09-654-617-370318 Sequence 370318,
45 17.2 78.2 278 27 US-09-684-016-370318 Sequence 370318,

ALIGNMENTS

RESULT 1
US-09-380-826A-2
; Sequence 2, Application US/09380826A
; GENERAL INFORMATION:
; APPLICANT: Chappell, Rod
; TITLE OF INVENTION: LEPTOSPIRA PATHOGENS
; FILE REFERENCE: DAVIE79.001APC
; CURRENT APPLICATION NUMBER: US/09/380,826A
; CURRENT FILING DATE: 1999-11-22
; PRIOR APPLICATION NUMBER: PCT/AU98/00145
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: AU P05494/97
; PRIOR FILING DATE: 1997-03-07
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Primer for L. fainei.
US-09-380-826A-2

Query Match 100.0%; Score 22; DB 17; Length 22;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 tttggtacacagaatttgata 22
|||
Db 1 tttggtacacagaatttgata 22
RESULT 2
US-09-380-826A-7
; Sequence 7, Application US/09380826A
; GENERAL INFORMATION:
; APPLICANT: Chappell, Rod
; TITLE OF INVENTION: LEPTOSPIRA PATHOGENS
; FILE REFERENCE: DAVIE79.001APC
; CURRENT APPLICATION NUMBER: US/09/380,826A
; CURRENT FILING DATE: 1999-11-22
; PRIOR APPLICATION NUMBER: PCT/AU98/00145
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: AU P05494/97
; PRIOR FILING DATE: 1997-03-07
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Primer for L. fainei.
US-09-380-826A-7

Query Match 100.0%; Score 22; DB 17; Length 22;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 tttggtacacagaatttgata 22
|||
Db 1 tttggtacacagaatttgata 22

RESULT 3
US-09-380-826A-1
; Sequence 1, Application US/09380826A
; GENERAL INFORMATION:
; APPLICANT: Chappell, Rod
; TITLE OF INVENTION: LEPTOSPIRA PATHOGENS
; FILE REFERENCE: DAVIE79.001APC
; CURRENT APPLICATION NUMBER: US/09/380,826A
; CURRENT FILING DATE: 1999-11-22
; PRIOR APPLICATION NUMBER: PCT/AU98/00145
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: AU P05494/97
; PRIOR FILING DATE: 1997-03-07
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1477
; TYPE: DNA
; ORGANISM: Leptospira fainei
US-09-380-826A-1

Query Match 100.0%; Score 22; DB 17; Length 1477;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 tttggtacacagaatttgata 22
|||
Db 154 tttggtacacagaatttgata 175

RESULT 4
US-60-132-861-11944
; Sequence 11944, Application US/60132861
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: ASPERGILLUS NIDULANS GENOME SEQUENCE AND USES THEREOF
; FILE REFERENCE: 38-21(15485)C
; CURRENT APPLICATION NUMBER: US/60/132,861
; CURRENT FILING DATE: 1999-05-07
; NUMBER OF SEQ ID NOS: 13800
; SEQ ID NO 11944
; LENGTH: 420
; TYPE: DNA
; ORGANISM: Aspergillus nidulans
US-60-132-861-11944

Query Match 85.5%; Score 18.8; DB 46; Length 420;
Best Local Similarity 90.9%; Pred. No. 1.1e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 tttggtacacagaatttgata 22
|||
Db 234 tttggtacacagaatttgata 255

RESULT 5
US-09-649-165A-7425/C
; Sequence 7425, Application US/09649165A
; GENERAL INFORMATION:
; APPLICANT: Guegler, Karl J.
; APPLICANT: Doyle, Martin
; APPLICANT: Momiyama, Monika
; TITLE OF INVENTION: DETECTING EXPRESSION OF ARABIDOPSIS THALIANA GENES
; FILE REFERENCE: PA-0019 US
; CURRENT APPLICATION NUMBER: US/09/649,165A
; CURRENT FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 60/150,517

;; PRIOR FILING DATE: 1999-08-24
;; NUMBER OF SEQ ID NOS: 7753
;; SOFTWARE: PERL Program
;; SEQ ID NO 7425
;; LENGTH: 530
;; TYPE: DNA
;; ORGANISM: Arabidopsis thaliana
;; FEATURE:
;; NAME/KEY: misc.feature
;; OTHER INFORMATION: Incyte ID No: 701676626
;; LOCATION: 6
;; NAME/KEY: unsure
;; OTHER INFORMATION: a, t, c, g, or other
US-09-649-165A-7425

Query Match 85.5%; Score 18.8; DB 25; Length 530;
Best Local Similarity 90.9%; Pred. No. 1.1e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 tgttgatcacagaattgata 22
||||| |||||||||
DB 254 tgttgatcacagaattgata 233

RESULT 6
US-09-654-617-175385
;; Sequence 175385, Application US/09654617
;; GENERAL INFORMATION:
;; APPLICANT: Kovacic, David K.
;; APPLICANT: Liu, Jingdong
;; TITLE OF INVENTION: Annotated Plant Genes
;; FILE REFERENCE: 38-21(15097)D
;; CURRENT APPLICATION NUMBER: US/09/654,617
;; CURRENT FILING DATE: 2000-09-05
;; NUMBER OF SEQ ID NOS: 463173
;; SEQ ID NO 175385
;; LENGTH: 581
;; TYPE: DNA
;; ORGANISM: Arabidopsis thaliana
US-09-654-617-175385

Query Match 85.5%; Score 18.8; DB 25; Length 581;
Best Local Similarity 90.9%; Pred. No. 1.1e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 tgttgatcacagaattgata 22
||||| |||||||||
DB 420 tgttgatcacagaattgata 441

RESULT 7
US-09-684-016-175385
;; Sequence 175385, Application US/09684016
;; GENERAL INFORMATION:
;; APPLICANT: Liu, Jingdong, David K.
;; APPLICANT: Kovacic, David K.
;; TITLE OF INVENTION: Annotated Plant Genes
;; FILE REFERENCE: 38-21(15097)D
;; CURRENT APPLICATION NUMBER: US/09/684,016
;; CURRENT FILING DATE: 2000-10-10
;; PRIOR APPLICATION NUMBER: US 09/654,617
;; PRIOR FILING DATE: 2000-09-05
;; NUMBER OF SEQ ID NOS: 463173
;; SEQ ID NO 175385
;; LENGTH: 581
;; TYPE: DNA
;; ORGANISM: Arabidopsis thaliana
US-09-684-016-175385

Query Match 85.5%; Score 18.8; DB 27; Length 581;

Best Local Similarity 90.9%; Pred. No. 1.1e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 tgttgatcacagaattgata 22
||||| |||||||||
DB 420 tgttgatcacagaattgata 441

RESULT 8
US-09-565-309A-20908
;; Sequence 20908, Application US/09565309A
;; GENERAL INFORMATION:
;; APPLICANT: ALEXANDROV, Nikolai
;; APPLICANT: BROVER, Vyacheslav
;; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
;; FILE REFERENCE: 2750-0853P
;; CURRENT APPLICATION NUMBER: US/09/565,309A
;; CURRENT FILING DATE: 2000-05-05
;; NUMBER OF SEQ ID NOS: 68449
;; SEQ ID NO 20908
;; LENGTH: 649
;; TYPE: DNA
;; ORGANISM: Arabidopsis thaliana
;; FEATURE:
;; NAME/KEY: misc.feature
;; LOCATION: (1)..(649)
;; OTHER INFORMATION: any n = a, g, c, t, unknown, or other
;; NAME/KEY: misc.feature
;; LOCATION: (1)..(649)
;; OTHER INFORMATION: 27496:110925 (Clone Number:Unique Sequence Identifier)
US-09-565-309A-20908

Query Match 85.5%; Score 18.8; DB 22; Length 649;
Best Local Similarity 90.9%; Pred. No. 1.1e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 tgttgatcacagaattgata 22
||||| |||||||||
DB 265 tgttgatcacagaattgata 286

RESULT 9
US-09-565-309A-57118
;; Sequence 57118, Application US/09565309A
;; GENERAL INFORMATION:
;; APPLICANT: ALEXANDROV, Nikolai
;; APPLICANT: BROVER, Vyacheslav
;; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
;; FILE REFERENCE: 2750-0853P
;; CURRENT APPLICATION NUMBER: US/09/565,309A
;; CURRENT FILING DATE: 2000-05-05
;; NUMBER OF SEQ ID NOS: 68449
;; SEQ ID NO 57118
;; LENGTH: 1069
;; TYPE: DNA
;; ORGANISM: Arabidopsis thaliana
;; FEATURE:
;; NAME/KEY: misc.feature
;; LOCATION: (1)..(1069)
;; OTHER INFORMATION: any n = a, g, c, t, unknown, or other
;; NAME/KEY: misc.feature
;; LOCATION: (1)..(1069)
;; OTHER INFORMATION: 27496 : OVERLAP (Clone Number : OVERLAP)
US-09-565-309A-57118

Query Match 85.5%; Score 18.8; DB 22; Length 1069;
Best Local Similarity 90.9%; Pred. No. 1.2e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 tttggtacacaagatttgata 22
||||| |||||||||
Db 690 tttggtacacaagatttgaca 711

RESULT 10

US-09-565-309A-49885
; Sequence 49885, Application US/09565309A
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai
; APPLICANT: BROVER, Vyacheslav
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: THEREBY
; FILE REFERENCE: 2750-0853P
; CURRENT APPLICATION NUMBER: US/09/565,309A
; CURRENT FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 68449
; SEQ ID NO 49885
; LENGTH: 1072
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: (1)..(1072)
; OTHER INFORMATION: any n = a, g, c, t, unknown, or other
; NAME/KEY: misc-feature
; LOCATION: (1)..(1072)
; OTHER INFORMATION: 27496 : 5TAG CONSENSUS (Clone Number:5tag_consensus)
US-09-565-309A-49885

Query Match 85.5%; Score 18.8; DB 22; Length 1072;
Best Local Similarity 90.9%; Pred. No. 1.2e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 tttggtacacaagatttgata 22
||||| |||||||||
Db 691 tttggtacacaagatttgaca 712

RESULT 11

US-09-654-617-117747
; Sequence 117747, Application US/09654617
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: LIU, Jindong
; TITLE OF INVENTION: Annotated Plant Genes
; FILE REFERENCE: 38-21(15097)D
; CURRENT APPLICATION NUMBER: US/09/654,617
; CURRENT FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 463173
; SEQ ID NO 117747
; LENGTH: 1128
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; OTHER INFORMATION: unsure at all n locations
US-09-654-617-117747

Query Match 85.5%; Score 18.8; DB 25; Length 1128;
Best Local Similarity 90.9%; Pred. No. 1.2e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 tttggtacacaagatttgata 22
||||| |||||||||
Db 719 tttggtacacaagatttgaca 740

RESULT 12

US-09-684-016-117747
; Sequence 117747, Application US/09684016
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.

APPLICANT: LIU, Jindong Annotated Plant Genes
; TITLE OF INVENTION: 38-21(15097)D
; FILE REFERENCE: 38-21(15097)D
; CURRENT APPLICATION NUMBER: US/09/684,016
; CURRENT FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: US 09/654,617
; PRIOR FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 463173
; SEQ ID NO 117747
; LENGTH: 1128
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1128)
; OTHER INFORMATION: unsure at all n locations
US-09-684-016-117747

Query Match 85.5%; Score 18.8; DB 27; Length 1128;
Best Local Similarity 90.9%; Pred. No. 1.2e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 tttggtacacaagatttgata 22
||||| |||||||||
Db 719 tttggtacacaagatttgaca 740

RESULT 13

US-60-138-103-11564/C
; Sequence 11564, Application US/60138103
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Timberlake, William E.
; TITLE OF INVENTION: ASPERGILLUS NIDULANS GENOME SEQUENCE AND USES THEREOF
; FILE REFERENCE: 38-10(15485)D
; CURRENT APPLICATION NUMBER: US/60/138,103
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 28006
; SEQ ID NO 11564
; LENGTH: 1429
; TYPE: DNA
; ORGANISM: Aspergillus nidulans
US-60-138-103-11564

Query Match 85.5%; Score 18.8; DB 46; Length 1429;
Best Local Similarity 90.9%; Pred. No. 1.2e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 tttggtacacaagatttgata 22
||||| |||||||||
Db 1206 tttggtacacaagatttgaca 1185

RESULT 14

US-09-404-520-5289
; Sequence 5289, Application US/09404520
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Ghodasli, Azita
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: McIninch, James
; APPLICANT: Timberlake, William E.
; APPLICANT: Yu, Jaehyuk
; TITLE OF INVENTION: EMERICELLA NIDULANS GENOME SEQUENCE AND USES THEREOF
; FILE REFERENCE: 38-10(15498)A
; CURRENT APPLICATION NUMBER: US/09/404,520
; CURRENT FILING DATE: 1999-09-23
; NUMBER OF SEQ ID NOS: 44345
; SEQ ID NO 5289
; LENGTH: 5874
; TYPE: DNA

ORGANISM: Aspergillus nidulans
US-09-404-520-5289

Query Match 85.5%; Score 18.8; DB 18; Length 5874;
Best Local Similarity 90.9%; Pred. No. 1.4e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 tgttgatcacagaatttgata 22
||||| ||||||| |||||
DB 248 tgttgatcacagaatttgata 269

RESULT 15
US-09-534-859-352

Sequence 352, Application US/09534859
GENERAL INFORMATION:
APPLICANT: Bush, David F.
APPLICANT: Last, Robert L.
APPLICANT: Levin, Irena M.
APPLICANT: Norris, Susan R.
APPLICANT: Parnell, Laurence D.
APPLICANT: Rounsley, Steven D.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: PLANT POLYMORPHIC MARKERS AND USES THEREOF
FILE REFERENCE: 38-10(15493)B
CURRENT APPLICATION NUMBER: US/09/534,859
CURRENT FILING DATE: 2000-03-29
NUMBER OF SEQ ID NOS: 1127
SEQ ID NO 352
LENGTH: 92624
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-534-859-352

Query Match 85.5%; Score 18.8; DB 20; Length 92624;
Best Local Similarity 90.9%; Pred. No. 2e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 tgttgatcacagaatttgata 22
||||| ||||||| |||||
DB 35807 tgttgatcacagaatttgata 35828

Search completed: December 15, 2001, 05:42:55
Job time: 15820 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 05:58:16 : Search time 474.83 Seconds

(without alignments)
66.629 Million cell updates/sec

Title: US-09-380-826a-2

Perfect score: 22

Sequence: 1 tttgtgacacaaagattgata 22

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 973175 seqs, 719035916 residues

Total number of hits satisfying chosen parameters: 1946350

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_NA_New:*
1: /cgn2_6/ptodata/1/pna/PCF_NEW_COMB.seq:*
2: /cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq:*
3: /cgn2_6/ptodata/1/pna/US07_NEW_COMB.seq:*
4: /cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq:*
5: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:*
6: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq1:*
7: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:*
8: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18.8	85.5	1069	6 US-09-620-394B-4668	Sequence 4668, Ap
2	17.2	78.2	249	5 US-09-924-035A-520	Sequence 520, App
3	17.2	78.2	622	6 US-09-985-678-211824	Sequence 211824, A
4	17.2	78.2	2038	6 US-09-922-279A-2773	Sequence 2773, Ap
5	17.2	78.2	46891	6 US-09-815-264-82805	Sequence 82805, A
6	16.8	76.4	267	6 US-09-982-402-1036	Sequence 1036, Ap
7	16.8	76.4	551	6 US-09-849-526A-27340	Sequence 27340, A
8	16.8	76.4	849	6 US-09-815-264-26466	Sequence 26466, A
9	16.8	76.4	1246	6 US-09-815-264-916319	Sequence 916319, A
10	16.8	76.4	1246	6 US-09-815-264-916319	Sequence 916319, A
11	16.8	76.4	1501	6 US-09-815-264-23140	Sequence 23140, A
12	16.8	76.4	1501	6 US-09-815-264-46882	Sequence 46882, A
13	16.8	76.4	1501	6 US-09-815-264-53250	Sequence 53250, A
14	16.8	76.4	6940	6 US-09-815-264-76889	Sequence 76889, A
15	16.8	76.4	12081	6 US-09-815-264-81045	Sequence 81045, A
16	16.8	76.4	12198	6 US-09-815-264-76653	Sequence 76653, A
17	16.8	76.4	41269	6 US-09-815-264-59808	Sequence 59808, A
18	16.8	76.4	50929	6 US-09-815-264-61033	Sequence 61033, A
19	16.8	76.4	67216	6 US-09-815-264-71944	Sequence 71944, A
20	16.4	74.5	360	6 US-09-388-906A-11536	Sequence 11536, A
21	16.4	74.5	4861	6 US-09-815-264-109466	Sequence 109466, A
22	16.2	73.6	278	6 US-09-985-678-279880	Sequence 279880, A
23	16.2	73.6	376	6 US-09-912-293-18469	Sequence 18469, A
24	16.2	73.6	377	6 US-09-845-487A-2238	Sequence 2238, App
25	16.2	73.6	419	6 US-09-933-524A-2094	Sequence 2094, Ap

ALIGNMENTS

26	16.2	73.6	461	6 US-09-849-526A-12135	Sequence 12135, A
27	16.2	73.6	845	6 US-09-922-279A-1268	Sequence 1268, Ap
28	16.2	73.6	1251	6 US-09-898-888A-9009	Sequence 9009, Ap
29	16.2	73.6	1457	6 US-09-815-264-60838	Sequence 60838, A
30	16.2	73.6	1501	6 US-09-815-264-39775	Sequence 39775, A
31	16.2	73.6	1501	6 US-09-815-264-39775	Sequence 43531, A
32	16.2	73.6	1501	6 US-09-815-264-43531	Sequence 43531, A
33	16.2	73.6	3603	6 US-09-815-264-80803	Sequence 80803, A
34	16.2	73.6	6237	6 US-09-815-264-72243	Sequence 72243, A
35	16.2	73.6	7856	6 US-09-815-264-86401	Sequence 86401, A
36	16.2	73.6	9723	6 US-09-815-264-75693	Sequence 75693, A
37	16.2	73.6	10839	6 US-09-815-264-78946	Sequence 78946, A
38	16.2	73.6	26601	6 US-09-815-264-74659	Sequence 74659, A
39	16.2	73.6	32005	6 US-09-815-264-73015	Sequence 73015, A
40	16.2	73.6	36227	6 US-09-815-264-66819	Sequence 66819, A
41	16	72.7	610	6 US-09-985-678-213635	Sequence 213635, A
42	16	72.7	1381	6 US-09-815-264-108378	Sequence 108378, A
43	16	72.7	37658	6 US-09-815-264-74259	Sequence 74259, A
44	16	72.7	39365	6 US-09-815-264-79204	Sequence 79204, A
45	15.8	71.8	270	6 US-09-985-678-42368	Sequence 42368, A

```

RESULT 1
US-09-620-394B-4668
; Sequence 4668, Application US/09620394B
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptid
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 2750-1067P
; CURRENT APPLICATION NUMBER: US/09/620,394B
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 9131
; SEQ ID NO 4668
; LENGTH: 1069
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..1069
; OTHER INFORMATION: any n = a, g, c, t, unknown, or other
; NAME/KEY: misc.feature
; LOCATION: 1..1069
; OTHER INFORMATION: Ceres Seq. ID 1393173
US-09-620-394B-4668

Query Match 85.5%; Score 18.8; DB 6; Length 1069;
Best Local Similarity 90.9%; Pred. No. 16;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 690 tttgtgacacaaagattgata 22
||||| |||||||||
1 tttgtgacacaaagattgata 22
||||| |||||||||
Db 690 tttgtgacacaaagattgata 711

RESULT 2
US-09-924-035A-520
; Sequence 520, Application US/09924035A
; GENERAL INFORMATION:
; APPLICANT: Grilach, Jrn
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; TITLE OF INVENTION: thaliana
; FILE REFERENCE: 2011US
; CURRENT APPLICATION NUMBER: US/09/924,035A
; CURRENT FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: US 60/148,784
; PRIOR FILING DATE: 1999-08-13
; NUMBER OF SEQ ID NOS: 900

```

```

; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 520
; LENGTH: 249
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(249)
; OTHER INFORMATION: n = A,T,C or G
US-09-924-035A-520

Query Match          78.2%; Score 17.2; DB 5; Length 249;
Best Local Similarity 86.4%; Pred. No. 81;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 tgttgatcacaaagtgtgata 22
||||| 1 ||||||| 11
Db 54 tgttgatcacaaagtgtgata 75

RESULT 3
US-09-985-678-211824
; Sequence 211824, Application US/09985678
; GENERAL INFORMATION:
; APPLICANT: Cheikh, Norline
; APPLICANT: Liu, Jingdong
; TITLE OF INVENTION: Annotated Plant Genes
; FILE REFERENCE: 16517.255/38-21(15097)F
; CURRENT APPLICATION NUMBER: US/09/985,678
; CURRENT FILING DATE: 2001-11-05
; PRIOR APPLICATION NUMBER: US 09/304,517
; PRIOR FILING DATE: 1999-05-06
; NUMBER OF SEQ ID NOS: 295529
; SEQ ID NO 211824
; LENGTH: 622
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)...(622)
; OTHER INFORMATION: unsure at all n locations
US-09-985-678-211824

Query Match          78.2%; Score 17.2; DB 6; Length 622;
Best Local Similarity 86.4%; Pred. No. 87;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 tgttgatcacaaagtgtgata 22
||||| 1 ||||||| 11
Db 520 tgttgatcacaaagtgtgata 541

RESULT 4
US-09-922-279A-2773
; Sequence 2773, Application US/09922279A
; GENERAL INFORMATION:
; APPLICANT: Tang, Yuanhua T.
; APPLICANT: Tillinghast, John
; APPLICANT: Slinku, Ankura
; APPLICANT: Liu, Chenghua
; APPLICANT: Dirmnac, Radoje T.
; TITLE OF INVENTION: Novel Contigs Obtained
; TITLE OF INVENTION: From Various Libraries
; FILE REFERENCE: 785
; CURRENT APPLICATION NUMBER: US/09/922,279A
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/491,404
; PRIOR FILING DATE: 2000-01-25
; NUMBER OF SEQ ID NOS: 3796
; SOFTWARE: PL_SP_genes Version 1.0
; SEQ ID NO 2773

```

```

; LENGTH: 2038
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: (1091)...(956)
; OTHER INFORMATION: this location contains the signal peptide sequence,
; OTHER INFORMATION: MYLLLPFGVYSIRSLFLGRPIQGITSDFTLLFENLDSMPLS, Run with Signal
; NAME/KEY: misc_feature
; LOCATION: (605)...(1902)
; OTHER INFORMATION: similar to g15931821 in the genepept database release 114,
; OTHER INFORMATION: Run with FASTXY 3.3f00, default parameters
US-09-922-279A-2773

Query Match          78.2%; Score 17.2; DB 6; Length 2038;
Best Local Similarity 86.4%; Pred. No. 96;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 tgttgatcacaaagtgtgata 22
||||| 1 ||||||| 11
Db 1929 tgttgatcacaaagtgtgata 1950

RESULT 5
US-09-815-264-82805
; Sequence 82805, Application US/09815264
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Cao, Yongwei
; APPLICANT: Dotson, Stanton B.
; APPLICANT: Koshi, Jeffrey M.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: McIninch, James
; APPLICANT: Wu, Mei
; TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof
; FILE REFERENCE: 38-21(5137)G
; CURRENT APPLICATION NUMBER: US/09/815,264
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 09/620,392
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/702,134
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 109669
; SEQ ID NO 82805
; LENGTH: 46891
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-815-264-82805

Query Match          78.2%; Score 17.2; DB 6; Length 46891;
Best Local Similarity 86.4%; Pred. No. 1,3e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 tgttgatcacaaagtgtgata 22
||||| 1 ||||||| 11
Db 45218 tgttgatcacaaagtgtgata 45239

RESULT 6
US-09-982-402-1036
; Sequence 1036, Application US/09982402
; GENERAL INFORMATION:
; APPLICANT: [list inventors here]
; TITLE OF INVENTION: [list title here]
; FILE REFERENCE: [list docket ID here]
; CURRENT APPLICATION NUMBER: US/09/982,402
; CURRENT FILING DATE: 2001-10-17
; NUMBER OF SEQ ID NOS: 3553
; SOFTWARE: PERL Program
; SEQ ID NO 1036

```

```

; LENGTH: 267
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 700196524H1
US-09-982-402-1036
```

```

Query Match          76.4%; Score 16.8; DB 6; Length 267;
Best Local Similarity 90.0%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 2 gttggatcacagaatttgat 21
    ||||| ||||| ||||| |||
DB 180 gttggatgacagaattcgat 199
```

```

RESULT 7
US-09-849-526A-27340/c
; Sequence 27340, Application US/09849526A
; GENERAL INFORMATION:
; APPLICANT: Ruff, Thomas G.
; APPLICANT: Shukla, Hridayabhimanjan
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 16517,250138-21(51930)B)
; CURRENT APPLICATION NUMBER: US/09/849,526A
; PRIOR FILING DATE: 2001-05-07
; PRIOR APPLICATION NUMBER: US 60/202,214
; PRIOR FILING DATE: 2000-05-08
; PRIOR APPLICATION NUMBER: US 09/654,617
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: US 09/684,016
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: US 09/816,660
; PRIOR FILING DATE: 2001-03-26
; NUMBER OF SEQ ID NOS: 30131
; SEQ ID NO 27340
; LENGTH: 551
; TYPE: DNA
; ORGANISM: Zea mays subsp. mexicana
; FEATURE:
; OTHER INFORMATION: Clone ID: uc-zmflteosinte90d04b1
US-09-849-526A-27340
```

```

Query Match          76.4%; Score 16.8; DB 6; Length 551;
Best Local Similarity 90.0%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 3 ttggatcacagaatttgata 22
    ||| | ||||| ||||| |||
DB 264 ttgcagcacaagatttgata 245
```

```

RESULT 8
US-09-815-264-26466/c
; Sequence 26466, Application US/09815264
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Cao, Yongwei
; APPLICANT: Dotson, Stanton B.
; APPLICANT: Koshi, Jeffrey M.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: McIninch, James
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51237)G
; CURRENT APPLICATION NUMBER: US/09/815,264
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 09/620,392
```

```

; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/702,134
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 109669
; SEQ ID NO 26466
; LENGTH: 849
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-815-264-26466
```

```

Query Match          76.4%; Score 16.8; DB 6; Length 849;
Best Local Similarity 90.0%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 2 gttggatcacagaatttgat 21
    ||||| ||||| ||||| |||
DB 256 gttggatcactagatttagat 237
```

```

RESULT 9
US-09-815-264-63319
; Sequence 63319, Application US/09815264
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Cao, Yongwei
; APPLICANT: Dotson, Stanton B.
; APPLICANT: Koshi, Jeffrey M.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: McIninch, James
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51237)G
; CURRENT APPLICATION NUMBER: US/09/815,264
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 09/620,392
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/702,134
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 109669
; SEQ ID NO 63319
; LENGTH: 1046
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-815-264-63319
```

```

Query Match          76.4%; Score 16.8; DB 6; Length 1046;
Best Local Similarity 90.0%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 2 gttggatcacagaatttgat 21
    ||||| ||||| ||||| |||
DB 791 gttggatcactagatttagat 810
```

```

RESULT 10
US-09-815-264-91672
; Sequence 91672, Application US/09815264
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Cao, Yongwei
; APPLICANT: Dotson, Stanton B.
; APPLICANT: Koshi, Jeffrey M.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: McIninch, James
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51237)G
; CURRENT APPLICATION NUMBER: US/09/815,264
; CURRENT FILING DATE: 2001-03-23
```

```

; PRIOR APPLICATION NUMBER: US 09/620,392
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/702,134
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 109669
; SEQ ID NO 91672
; LENGTH: 1246
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-815-264-91672

Query Match          76.4%; Score 16.8; DB 6; Length 1246;
Best Local Similarity 90.0%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 2 gttgatacacaagatttgat 21
    ||||| ||||| ||||| |||
Db 1013 gttgatacactagattagat 1032

RESULT 11
US-09-815-264-23140
; Sequence 23140, Application US/09815264
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Cao, Yongwei
; APPLICANT: Dotson, Stanton B.
; APPLICANT: Koshi, Jeffrey M.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jindong
; APPLICANT: McIninch, James
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51237)G
; CURRENT APPLICATION NUMBER: US 09/815,264
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 09/620,392
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/702,134
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 109669
; SEQ ID NO 23140
; LENGTH: 1501
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-815-264-23140

Query Match          76.4%; Score 16.8; DB 6; Length 1501;
Best Local Similarity 90.0%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 3 ttgatacacaagatttgata 22
    ||||| || ||||| ||||| |||
Db 674 ttgatacacaagatttgata 693

RESULT 12
US-09-815-264-46882
; Sequence 46882, Application US/09815264
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Cao, Yongwei
; APPLICANT: Dotson, Stanton B.
; APPLICANT: Koshi, Jeffrey M.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jindong
; APPLICANT: McIninch, James
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51237)G
; CURRENT APPLICATION NUMBER: US 09/815,264
```

```

; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 09/620,392
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/702,134
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 109669
; SEQ ID NO 46882
; LENGTH: 1501
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1501)
; OTHER INFORMATION: unsure at all n locations
US-09-815-264-46882
```

```

Query Match          76.4%; Score 16.8; DB 6; Length 1501;
Best Local Similarity 90.0%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 3 ttgatacacaagatttgata 22
    ||||| || ||||| ||||| |||
Db 674 ttgatacacaagatttgata 693
```

```

RESULT 13
US-09-815-264-53250
; Sequence 53250, Application US/09815264
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Cao, Yongwei
; APPLICANT: Dotson, Stanton B.
; APPLICANT: Koshi, Jeffrey M.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jindong
; APPLICANT: McIninch, James
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51237)G
; CURRENT APPLICATION NUMBER: US 09/815,264
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 09/620,392
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/702,134
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 109669
; SEQ ID NO 53250
; LENGTH: 1501
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-815-264-53250
```

```

Query Match          76.4%; Score 16.8; DB 6; Length 1501;
Best Local Similarity 90.0%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 2 gttgatacacaagatttgat 21
    ||||| ||||| ||||| |||
Db 88 gttgatacactagattagat 107
```

```

RESULT 14
US-09-815-264-76889
; Sequence 76889, Application US/09815264
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Cao, Yongwei
; APPLICANT: Dotson, Stanton B.
; APPLICANT: Koshi, Jeffrey M.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jindong
```

```

; APPLICANT: McIninch, James
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51237)G
; CURRENT APPLICATION NUMBER: US/09/815,264
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 09/620,392
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/702,134
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 109669
; SEQ ID NO 76889
; LENGTH: 6940
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(6940)
; OTHER INFORMATION: unsure at all n locations
US-09-815-264-76889
```

```

Query Match          76.4%; Score 16.8; DB 6; Length 6940;
Best Local Similarity 90.0%; Pred. No. 1.7e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      3 ttggatcacaaatttgata 22
      ||||| || ||||| |||||
DB      3165 ttggaccacaaagatttgata 3184
```

```

RESULT 15
US-09-815-264-81045/C
; Sequence 81045, Application US/09815264
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Cao, Yongwei
; APPLICANT: Dotson, Stanton B.
; APPLICANT: Koshl, Jeffrey M.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: McIninch, James
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51237)G
; CURRENT APPLICATION NUMBER: US/09/815,264
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 09/620,392
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/702,134
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 109669
; SEQ ID NO 81045
; LENGTH: 12081
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-815-264-81045
```

```

Query Match          76.4%; Score 16.8; DB 6; Length 12081;
Best Local Similarity 90.0%; Pred. No. 1.7e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      2 gttgatacaaaatttgat 21
      ||||| ||||| ||||| |||
DB      4391 GTTGATCACTAGATTAGAT 4372
```

Search completed: December 15, 2001, 05:58:19
Job time: 12262 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 05:42:55 ; Search time 7962.77 seconds
(without alignments)
15.407 Million cell updates/sec

Title: US-09-380-826a-4
Perfect score: 1 t9t9ga 7
Sequence: 1 t9t9ga 7

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 17159718 seqs, 876320856 residues

Total number of hits satisfying chosen parameters: 34319436

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%

Listing first 45 summaries

Database : Pending_Patents_Nr_Main:*

```
1: /cgn2_6/pdata/2/pna/US0911.COMB.seq:*
2: /cgn2_6/pdata/2/pna/US0612.COMB.seq:*
3: /cgn2_6/pdata/2/pna/US0613.COMB.seq:*
4: /cgn2_6/pdata/2/pna/US0614.COMB.seq:*
5: /cgn2_6/pdata/2/pna/US0615.COMB.seq:*
6: /cgn2_6/pdata/2/pna/US0616.COMB.seq:*
7: /cgn2_6/pdata/2/pna/US0617.COMB.seq:*
8: /cgn2_6/pdata/2/pna/US0618.COMB.seq:*
9: /cgn2_6/pdata/2/pna/US0619.COMB.seq:*
10: /cgn2_6/pdata/2/pna/US0620.COMB.seq:*
11: /cgn2_6/pdata/2/pna/US0621.COMB.seq:*
12: /cgn2_6/pdata/2/pna/US0622.COMB.seq:*
13: /cgn2_6/pdata/2/pna/US0623.COMB.seq:*
14: /cgn2_6/pdata/2/pna/US0624.COMB.seq:*
15: /cgn2_6/pdata/2/pna/US0625.COMB.seq:*
16: /cgn2_6/pdata/2/pna/US0626.COMB.seq:*
17: /cgn2_6/pdata/2/pna/US0627.COMB.seq:*
18: /cgn2_6/pdata/2/pna/US0628.COMB.seq:*
19: /cgn2_6/pdata/2/pna/US0629.COMB.seq:*
20: /cgn2_6/pdata/2/pna/US0630.COMB.seq:*
21: /cgn2_6/pdata/2/pna/US0631.COMB.seq:*
22: /cgn2_6/pdata/2/pna/US0632.COMB.seq:*
23: /cgn2_6/pdata/2/pna/US0633.COMB.seq:*
24: /cgn2_6/pdata/2/pna/US0634.COMB.seq:*
25: /cgn2_6/pdata/2/pna/US0635.COMB.seq:*
26: /cgn2_6/pdata/2/pna/US0636.COMB.seq:*
27: /cgn2_6/pdata/2/pna/US0637.COMB.seq:*
28: /cgn2_6/pdata/2/pna/US0638.COMB.seq:*
29: /cgn2_6/pdata/2/pna/US0639.COMB.seq:*
30: /cgn2_6/pdata/2/pna/US0640.COMB.seq:*
31: /cgn2_6/pdata/2/pna/US0641.COMB.seq:*
32: /cgn2_6/pdata/2/pna/US0642.COMB.seq:*
33: /cgn2_6/pdata/2/pna/US0643.COMB.seq:*
34: /cgn2_6/pdata/2/pna/US0644.COMB.seq:*
35: /cgn2_6/pdata/2/pna/US0645.COMB.seq:*
36: /cgn2_6/pdata/2/pna/US0646.COMB.seq:*
37: /cgn2_6/pdata/2/pna/US0647.COMB.seq:*
38: /cgn2_6/pdata/2/pna/US0648.COMB.seq:*
39: /cgn2_6/pdata/2/pna/US0649.COMB.seq:*
40: /cgn2_6/pdata/2/pna/US0650.COMB.seq:*
41: /cgn2_6/pdata/2/pna/US0651.COMB.seq:*
42: /cgn2_6/pdata/2/pna/US0652.COMB.seq:*
43: /cgn2_6/pdata/2/pna/US0653.COMB.seq:*
```

```
44: /cgn2_6/pdata/2/pna/US0654.COMB.seq:*
45: /cgn2_6/pdata/2/pna/US0655.COMB.seq:*
46: /cgn2_6/pdata/2/pna/US0656.COMB.seq:*
47: /cgn2_6/pdata/2/pna/US0657.COMB.seq:*
48: /cgn2_6/pdata/2/pna/US0658.COMB.seq:*
49: /cgn2_6/pdata/2/pna/US0659.COMB.seq:*
50: /cgn2_6/pdata/2/pna/US0660.COMB.seq:*
51: /cgn2_6/pdata/2/pna/US0661.COMB.seq:*
52: /cgn2_6/pdata/2/pna/US0662.COMB.seq:*
53: /cgn2_6/pdata/2/pna/US0663.COMB.seq:*
54: /cgn2_6/pdata/2/pna/US0664.COMB.seq:*
55: /cgn2_6/pdata/2/pna/US0665.COMB.seq:*
56: /cgn2_6/pdata/2/pna/US0666.COMB.seq:*
57: /cgn2_6/pdata/2/pna/US0667.COMB.seq:*
58: /cgn2_6/pdata/2/pna/US0668.COMB.seq:*
59: /cgn2_6/pdata/2/pna/US0669.COMB.seq:*
60: /cgn2_6/pdata/2/pna/US0670.COMB.seq:*
61: /cgn2_6/pdata/2/pna/US0671.COMB.seq:*
62: /cgn2_6/pdata/2/pna/US0672.COMB.seq:*
63: /cgn2_6/pdata/2/pna/US0673.COMB.seq:*
64: /cgn2_6/pdata/2/pna/US0674.COMB.seq:*
65: /cgn2_6/pdata/2/pna/US0675.COMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7	100.0	7	17 US-09-380-826a-4	Sequence 44, Appl
2	7	100.0	9	14 US-09-065-775-44	Sequence 44, Appl
3	7	100.0	32	US-09-916-466-170	Sequence 170, App
4	7	100.0	10	1 PCT-US01-24784-149	Sequence 149, App
5	7	100.0	10	1 PCT-US01-03620-47	Sequence 47, Appl
6	7	100.0	10	1 PCT-US01-18811-45	Sequence 45, Appl
7	7	100.0	10	1 PCT-US94-12882-13	Sequence 13, Appl
8	7	100.0	10	1 PCT-US99-13800-679	Sequence 679, App
9	7	100.0	10	5 US-08-141-969-13	Sequence 13, Appl
10	7	100.0	10	16 US-09-213-932-113	Sequence 13, Appl
11	7	100.0	10	17 US-09-335-032-2036	Sequence 2036, App
12	7	100.0	10	17 US-09-335-032-8949	Sequence 8949, App
13	7	100.0	10	17 US-09-335-032-11545	Sequence 11545, A
14	7	100.0	10	17 US-09-336-376-1209	Sequence 1209, App
15	7	100.0	10	17 US-09-336-376-4002	Sequence 4002, App
16	7	100.0	10	17 US-09-336-376-4328	Sequence 4328, App
17	7	100.0	10	17 US-09-336-376-5236	Sequence 5236, App
18	7	100.0	10	20 US-09-535-088-629	Sequence 629, Appl
19	7	100.0	10	23 US-09-601-844-18	Sequence 18, Appl
20	7	100.0	10	32 US-09-916-466-128	Sequence 128, Appl
21	7	100.0	10	32 US-09-945-505-30	Sequence 30, Appl
22	7	100.0	11	32 US-09-916-466-96	Sequence 96, Appl
23	7	100.0	12	1 PCT-US98-26935-386	Sequence 386, App
24	7	100.0	12	16 US-09-215-436-386	Sequence 386, App
25	7	100.0	13	19 US-09-528-209A-2234	Sequence 2234, App
26	7	100.0	13	19 US-09-528-209A-3083	Sequence 3083, App
27	7	100.0	13	19 US-09-528-209A-7388	Sequence 7388, App
28	7	100.0	12	32 US-09-586-479-2	Sequence 2, Appl
29	7	100.0	12	32 US-09-586-479-87	Sequence 87, Appl
30	7	100.0	13	1 PCT-US98-26935-274	Sequence 274, App
31	7	100.0	13	16 US-09-215-436-274	Sequence 274, App
32	7	100.0	13	19 US-09-528-209A-11570	Sequence 11570, App
33	7	100.0	14	15 US-09-164-961-1123	Sequence 1123, App
34	7	100.0	14	16 US-09-274-553-2956	Sequence 2956, App
35	7	100.0	14	16 US-09-274-553B-1407	Sequence 1407, App
36	7	100.0	14	16 US-09-274-553C-1407	Sequence 1407, App
37	7	100.0	14	16 US-09-277-026B-4887	Sequence 4887, App

	42	43	44	45	7	100.0	14	17	US-09-341-700A-419	Sequence 419, App
c	42	7	100.0	14	17	US-09-341-700A-421	Sequence 421, App			
c	43	7	100.0	14	17	US-09-341-700A-421	Sequence 421, App			
c	44	7	100.0	14	17	US-09-341-700A-421	Sequence 421, App			
c	45	7	100.0	14	19	US-09-504-231A-1407	Sequence 1407, Ap			

ALIGNMENTS

RESULT 1
US-09-380-826a-4
; Sequence 4, Application US/09380826A
; GENERAL INFORMATION:
; APPLICANT: Chappel, Rod
; TITLE OF INVENTION: LEPTOSPIRA PATHOGENS
; FILE REFERENCE: DAVIE79.001APC
; CURRENT APPLICATION NUMBER: US/09/380, 826A
; CURRENT FILING DATE: 1999-11-22
; PRIOR APPLICATION NUMBER: PCT/AU98/00145
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: AU P05494/97
; PRIOR FILING DATE: 1997-03-07
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 7
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Primer for L. fauvel.
US-09-380-826a-4

Query Match
Best Local Similarity 100.0%; Score 7; DB 17; Length 7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 tgttga 7
| | | | |
Db 1 tgttga 7

RESULT 2
US-09-065-775-44
; Sequence 44, Application US/09065775A
; GENERAL INFORMATION:
; APPLICANT: Cen, Hui
; TITLE OF INVENTION: Function-Based Gene Discovery
; FILE REFERENCE: 59818
; CURRENT APPLICATION NUMBER: US/09/065, 775A
; CURRENT FILING DATE: 1998-04-24
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 44
; LENGTH: 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: oligonucleotide primers
US-09-065-775-44

Query Match
Best Local Similarity 100.0%; Score 7; DB 14; Length 9;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 tgttga 7
| | | | |
Db 3 tgttga 9

RESULT 3
US-09-916-466-170
; Sequence 170, Application US/09916466
; GENERAL INFORMATION:
; APPLICANT: Ribozyne Pharmaceuticals, Inc.
; APPLICANT: Akhtar, Saghir
; APPLICANT: McSwiggen, Jim
; TITLE OF INVENTION: Enzymatic Nucleic Acid Treatment of Diseases or conditions Rel
; FILE REFERENCE: MBH00-958-J (400/032)
; CURRENT APPLICATION NUMBER: US/09/916, 466
; CURRENT FILING DATE: 2001-07-25
; NUMBER OF SEQ ID NOS: 446
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 170
; LENGTH: 9
; TYPE: RNA
; ORGANISM: Homo sapiens
US-09-916-466-170

Query Match
Best Local Similarity 57.1%; Score 7; DB 32; Length 9;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 tgttga 7
| | | | |
Db 1 ugnugga 7

RESULT 4
PCT-US00-24784-149
; Sequence 149, Application PC/TUS0024784
; GENERAL INFORMATION:
; APPLICANT: Nandabalan, Krishnan
; APPLICANT: Stephens, Joel Claiborne
; APPLICANT: Denton, R. Rex
; APPLICANT: Chew, Anne
; APPLICANT: Choi, Julie Y.
; TITLE OF INVENTION: DRUG TARGET ISOGENES;
; TITLE OF INVENTION: POLYMORPHISMS IN THE UNCOUPLING PROTEIN 3
; FILE REFERENCE: UCP3-0009PCT
; CURRENT APPLICATION NUMBER: PCT/US00/24784
; CURRENT FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: MMH009
; PRIOR FILING DATE: 1999-09-08
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 149
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US00-24784-149

Query Match
Best Local Similarity 100.0%; Score 7; DB 1; Length 10;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 tgttga 7
| | | | |
Db 1 tgttga 7

RESULT 5
PCT-US01-03620-47
; Sequence 47, Application PC/TUS0103620
; GENERAL INFORMATION:
; APPLICANT: Genaisance Pharmaceuticals, Inc.
; APPLICANT: Chew, Anne
; APPLICANT: Koshy, Beena
; APPLICANT: Choi, Julie Y.

APPLICANT: Stephens, J. Claiborne
TITLE OF INVENTION: DRUG TARGET ISOGENES: POLYMORPHISMS IN THE ANGIOTENSIN
RECEPTOR 2 GENE
FILE REFERENCE: MMH-0045 AGTR2
CURRENT APPLICATION NUMBER: PCT/US01/03620
CURRENT FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/179,922
PRIOR FILING DATE: 2000-02-03
NUMBER OF SEQ ID NOS: 74
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 47
LENGTH: 10
TYPE: DNA
ORGANISM: Homo sapiens
PCT-US01-03620-47

Query Match 100.0%; Score 7; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.9e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7
|111111|
Db 2 tgttga 8

RESULT 6
PCT-US01-18811A-45
Sequence 45, Application PC/TUS0118811A
GENERAL INFORMATION:
APPLICANT: Genesance Pharmaceuticals, Inc.
APPLICANT: Bentivegna, Steven C.
APPLICANT: Bieglecki, Karyn
APPLICANT: Duda, Amy
APPLICANT: Kazeml, Amir
APPLICANT: Koshy, Beena
TITLE OF INVENTION: Haplotypes of the BDKRB2 Gene
FILE REFERENCE: MMH-0740PCT BDKRB2
CURRENT APPLICATION NUMBER: PCT/US01/18811A
CURRENT FILING DATE: 2001-06-11
PRIOR APPLICATION NUMBER: 60/210,575
PRIOR FILING DATE: 2000-06-09
NUMBER OF SEQ ID NOS: 57
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 45
LENGTH: 10
TYPE: DNA
ORGANISM: Homo sapiens
PCT-US01-18811A-45

Query Match 100.0%; Score 7; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.9e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7
|111111|
Db 3 tgttga 9

RESULT 7
PCT-US94-12282-13/C
Sequence 13, Application PC/TUS9412282
GENERAL INFORMATION:
APPLICANT: Beattie, Kenneth L.
TITLE OF INVENTION: Microfabricated, Flowthrough Porous
Apparatus for Discrete Detection of Binding Reactions
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESS: Vinson & Elkins
STREET: 1455 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.

COUNTRY: U.S.A.
ZIP: 20004-1008
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/12282
FILING DATE: 27-Oct-1994
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/141,969
FILING DATE: 28-Oct-1994
ATTORNEY/AGENT INFORMATION:
NAME: Sanzo, Michael A.
REGISTRATION NUMBER: 36,912
REFERENCE/DOCKET NUMBER: HARC0001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)639-6500
TELEFAX: (202)639-6604
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
PCT-US94-12282-13

Query Match 100.0%; Score 7; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.9e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7
|111111|
Db 7 tgttga 1

RESULT 8
PCT-US99-13800-679
Sequence 679, Application PC/TUS9913800
GENERAL INFORMATION:
APPLICANT: Genzyme Corporation
APPLICANT: Roberts, Bruce L.
APPLICANT: Shankar, Srinivas
TITLE OF INVENTION: PREPARATION AND USE OF SUPERIOR VACCINES
FILE REFERENCE: 68126881206940
CURRENT APPLICATION NUMBER: PCT/US99/13800
CURRENT FILING DATE: 1999-06-18
EARLIER APPLICATION NUMBER: 60/090,039
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,040
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,041
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/089,853
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/089,997
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,079
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,035
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/089,993
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/089,992
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,072
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/089,878
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/089,991

EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,000
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,048
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/089,999
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,043
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,042
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,036
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,044
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/089,844
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,080
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/089,833
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/089,994
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,077
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,078
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,047
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,076
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,045
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/111,715
EARLIER FILING DATE: 1998-12-08
NUMBER OF SEQ ID NOS: 2138
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO: 679
LENGTH: 10
TYPE: DNA
ORGANISM: Homo sapiens
PCT-US99-13800-679

Query Match 100.0%; Score 7; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.9e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tgttga 7
1111111
DB 3 tgttga 9

RESULT 9
US-08-141-969-13/C
Sequence 13, Application US/08141969
GENERAL INFORMATION:
APPLICANT: Beattie, Kenneth L.
TITLE OF INVENTION: Microfabricated, Flowthrough Porous
TITLE OF INVENTION: Apparatus for Discrete Detection of Binding Reactions
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Vinson & Elkins
STREET: 1455 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20004-1008
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/141,969
FILING DATE: 28-OCT-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sanzo, Michael A.
REGISTRATION NUMBER: 36,912
REFERENCE/DOCKET NUMBER: HARC0001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)639-6500
TELEFAX: (202)639-6604
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
US-08-141-969-13

Query Match 100.0%; Score 7; DB 5; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.9e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tgttga 7
1111111
DB 7 TGTGGA 1

RESULT 10
US-09-213-932A-13/C
Sequence 13, Application US/09213932A
GENERAL INFORMATION:
APPLICANT: Beattie, Kenneth L.
TITLE OF INVENTION: Microfabricated, Flowthrough Porous
TITLE OF INVENTION: Apparatus for Discrete Detection of Binding Reactions
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Vinson & Elkins
STREET: 1455 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20004-1008
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/213,932A
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/631,751
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Sanzo, Michael A.
REGISTRATION NUMBER: 36,912
REFERENCE/DOCKET NUMBER: HARC0001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)639-6500
TELEFAX: (202)639-6604
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
US-09-213-932A-13

Query Match 100.0%; Score 7; DB 16; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.9e+05;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7
|111111
Db 7 TGTGGA 1

RESULT 11
US-09-335-032-2036/c
; Sequence 2036, Application US/09335032
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the Yeast
; FILE REFERENCE: 01107.78572
; CURRENT APPLICATION NUMBER: US/09/335,032
; CURRENT FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2036
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-335-032-2036

Query Match 100.0%; Score 7; DB 17; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.9e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7
|111111
Db 10 TGTGGA 4

RESULT 12
US-09-335-032-8949/c
; Sequence 8949, Application US/09335032
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the Yeast
; FILE REFERENCE: 01107.78572
; CURRENT APPLICATION NUMBER: US/09/335,032
; CURRENT FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8949
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-335-032-8949

Query Match 100.0%; Score 7; DB 17; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.9e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7
|111111
Db 10 TGTGGA 4

RESULT 13
US-09-335-032-11545
; Sequence 11545, Application US/09335032
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the Yeast
; FILE REFERENCE: 01107.78572
; CURRENT APPLICATION NUMBER: US/09/335,032
; CURRENT FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11545
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-335-032-11545

Query Match 100.0%; Score 7; DB 17; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.9e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7
|111111
Db 1 tgttga 7

RESULT 14
US-09-336-376-1209
; Sequence 1209, Application US/09336376
; GENERAL INFORMATION:
; APPLICANT: Roberts, Bruce L.
; TITLE OF INVENTION: POLYNUCLEOTIDE POPULATION ISOLATED FROM
; FILE REFERENCE: 126745205600
; CURRENT APPLICATION NUMBER: US/09/336,376
; CURRENT FILING DATE: 1999-06-17
; EARLIER APPLICATION NUMBER: 60/090,039
; EARLIER FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/090,040
; EARLIER FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/090,041
; EARLIER FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/089,853
; EARLIER FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/089,997
; EARLIER FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/090,079
; EARLIER FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/090,035
; EARLIER FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/089,993
; EARLIER FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/089,992
; EARLIER FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/090,072
; EARLIER FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/089,878
; EARLIER FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/089,991
; EARLIER FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/090,000
; EARLIER FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/090,048

```
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/089,999
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,043
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,042
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,036
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,044
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/089,844
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,080
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/089,833
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/089,994
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,077
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,078
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,047
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,076
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,045
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/111,715
EARLIER FILING DATE: 1998-12-08
NUMBER OF SEQ ID NOS: 5980
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO: 1209
LENGTH: 10
TYPE: DNA
ORGANISM: Homo sapiens
US-09-336-376-1209
```

```
Query Match      100.0%; Score 7; DB 17; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.9e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 tgttga 7
        |||
Db      4 tgttga 10

RESULT 15
US-09-336-376-4002
; Sequence 4002, Application us/09336376
; GENERAL INFORMATION:
; APPLICANT: Roberts, Bruce L.
; TITLE OF INVENTION: POLYNUCLEOTIDE POPULATION ISOLATED FROM
; FILE REFERENCE: 126745205600
; CURRENT APPLICATION NUMBER: us/09/336,376
; EARLIER FILING DATE: 1999-06-17
; EARLIER APPLICATION NUMBER: 60/090,039
; EARLIER FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/090,040
; EARLIER FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/090,041
; EARLIER FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/089,853
; EARLIER FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/089,997
; EARLIER FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/090,079
; EARLIER FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/090,035
; EARLIER FILING DATE: 1998-06-19
```

```
EARLIER APPLICATION NUMBER: 60/089,993
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/089,992
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,072
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/089,878
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/089,991
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,000
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,048
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/089,999
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,043
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,042
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,036
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,044
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/089,844
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,080
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/089,833
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/089,994
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,077
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,078
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,047
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,076
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/090,045
EARLIER FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/111,715
EARLIER FILING DATE: 1998-12-08
NUMBER OF SEQ ID NOS: 5980
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO: 4002
LENGTH: 10
TYPE: DNA
ORGANISM: Homo sapiens
US-09-336-376-4002
```

```
Query Match      100.0%; Score 7; DB 17; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.9e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 tgttga 7
        |||
Db      4 tgttga 10
```

Search completed: December 15, 2001, 05:42:56
Job time: 15821 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 05:58:19 ; Search time 474.83 Seconds
(without alignments)
21.200 Million cell updates/sec

```
Title: US-09-380-8
Perfect score: 7
Sequence: 1 tgttga 7
```

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 973175 seqs, 719035916 residues

Total number of hits satisfying chosen parameters: 1946350

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Post-processing:  Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

```
Database : Pending_Patients_NA_New:*
1 : /cgn2_6/ptodata/1/pna/pCT_NEM_COMB.seq:*
2 : /cgn2_6/ptodata/1/pna/US06_NEM_COMB.seq:*
3 : /cgn2_6/ptodata/1/pna/US07_NEM_COMB.seq:*
4 : /cgn2_6/ptodata/1/pna/US08_NEM_COMB.seq:*
5 : /cgn2_6/ptodata/1/pna/US09_NEM_COMB.seq:*
6 : /cgn2_6/ptodata/1/pna/US10_NEM_COMB.seq:1
7 : /cgn2_6/ptodata/1/pna/US10_NEM_COMB.seq:*
8 : /cgn2_6/ptodata/1/pna/US10_NEM_COMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	7	100.0	10	1	PCT-US01-30411-29	Sequence 29, Apple
C 2	7	100.0	13	6	US-09-825-805-181	Sequence 181, Apple
C 3	7	100.0	15	1	PCT-US01-30411-7	Sequence 7, Apple
C 4	7	100.0	15	1	PCT-US01-30411-17	Sequence 17, Apple
C 5	7	100.0	15	6	US-09-825-805-182	Sequence 182, Apple
C 6	7	100.0	16	6	US-09-527-030C-105	Sequence 105, Apple
C 7	7	100.0	17	6	US-09-527-030C-181	Sequence 181, Apple
C 8	7	100.0	17	6	US-09-685-664B-1636	Sequence 1636, Apple
C 9	7	100.0	17	6	US-09-685-664B-1637	Sequence 1637, Apple
C 10	7	100.0	17	6	US-09-685-664B-3484	Sequence 3484, Apple
C 11	7	100.0	17	6	US-09-685-664B-3485	Sequence 3485, Apple
C 12	7	100.0	17	6	US-09-685-664B-3486	Sequence 3486, Apple
C 13	7	100.0	17	6	US-09-685-664B-3493	Sequence 3493, Apple
C 14	7	100.0	17	6	US-09-740-332-1100	Sequence 1100, Apple
C 15	7	100.0	17	6	US-09-740-332-1101	Sequence 1101, Apple
C 16	7	100.0	17	6	US-09-740-332-1102	Sequence 1102, Apple
C 17	7	100.0	17	6	US-09-740-332-3453	Sequence 3453, Apple
C 18	7	100.0	17	6	US-09-740-332-3454	Sequence 3454, Apple
C 19	7	100.0	17	6	US-09-740-332-3455	Sequence 3455, Apple
C 20	7	100.0	17	6	US-09-870-161-1636	Sequence 1636, Apple
C 21	7	100.0	17	6	US-09-870-161-1637	Sequence 1637, Apple
C 22	7	100.0	17	6	US-09-870-161-3484	Sequence 3484, Apple
C 23	7	100.0	17	6	US-09-870-161-3485	Sequence 3485, Apple
C 24	7	100.0	17	6	US-09-870-161-3486	Sequence 3486, Apple
C 25	7	100.0	17	6	US-09-870-161-3493	Sequence 3493, Apple

ALIGNMENTS

26	100.0	17	6	US-09-870-161-6245	Sequence 6245, App
27	100.0	17	6	US-09-870-161-8559	Sequence 8559, App
28	7 100.0	17	6	US-09-730-289-116	Sequence 116, App
29	7 100.0	17	6	US-09-730-289-117	Sequence 117, App
30	7 100.0	17	6	US-09-730-289-118	Sequence 118, App
31	7 100.0	17	6	US-09-730-289-119	Sequence 762, App
32	7 100.0	17	6	US-09-730-289-962	Sequence 934, App
33	7 100.0	17	6	US-09-730-289-964	Sequence 1084, App
34	7 100.0	17	6	US-09-730-289-1085	Sequence 1085, App
35	7 100.0	17	6	US-09-776-474-70	Sequence 70, App1
36	7 100.0	17	6	US-09-776-474-71	Sequence 71, App1
37	7 100.0	17	6	US-09-776-474-72	Sequence 72, App1
38	7 100.0	17	6	US-09-776-474-425	Sequence 425, App
39	7 100.0	17	6	US-09-776-474-708	Sequence 708, App
40	7 100.0	17	6	US-09-776-474-709	Sequence 709, App
41	7 100.0	17	6	US-09-776-474-921	Sequence 921, App
42	7 100.0	17	6	US-09-776-474-1082	Sequence 1082, App
43	7 100.0	17	6	US-09-776-474-1083	Sequence 1083, App
44	7 100.0	17	6	US-09-825-000-539	Sequence 539, App
45	7 100.0	17	8	US-60-329-000-1311	Sequence 1311, App

```

RESULT 1
PCT-US01-30411-29/C
: Sequence 29, Application PC/TUS0130411
: GENERAL INFORMATION:
: APPLICANT: Genaisance Pharmaceuticals, Inc.
: APPLICANT: Choi, Julie Y
: APPLICANT: Klem, Stefanie E
: APPLICANT: Koshy, Beena
: APPLICANT: Parks, Katie E
: APPLICANT: Stephens, J. Claiborne
: TITLE OF INVENTION: HAPLOTYPES OF THE TCF2 GENE
: FILE REFERENCE: TCF2_MMH-1439PCT
: CURRENT APPLICATION NUMBER: PCT/US01/30411
: CURRENT FILING DATE: 2001-09-27
: PRIOR APPLICATION NUMBER: 60/235,710
: PRIOR FILING DATE: 2000-09-27
: NUMBER OF SEQ ID NOS: 36
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 29
: LENGTH: 10
: TYPE: DNA
: ORGANISM: Homo sapiens
PCT-US01-30411-29

Query Match          100.0%; Score 7; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 6,1e+04;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 tcttga 7
        |||||
Db       10 TGTGCA 4

RESULT 2
US-09-825-805-181
: Sequence 181, Application US/09825805
: GENERAL INFORMATION:
: APPLICANT: Ribozyme Pharmaceuticals, Inc.
: APPLICANT: Beigelman, Leo
: APPLICANT: Beaudry, Amber
: APPLICANT: Karpeisky, Alex
: APPLICANT: Adamic, Jasenka Matulic
: APPLICANT: Sweedler, Dave
: APPLICANT: Zinnen, Shawn
: TITLE OF INVENTION: Nucleotide Triphosphate and their Incorporation into Oligonucle
: FILE REFERENCE: MBH80-831-F (400/009)
: CURRENT APPLICATION NUMBER: US/09/825,805

```

```

; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 09/578,223
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 09/476,387
; PRIOR FILING DATE: 1999-12-30
; PRIOR APPLICATION NUMBER: 09/474,432
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 09/301,511
; PRIOR FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: 09/186,675
; PRIOR FILING DATE: 1998-11-04
; PRIOR APPLICATION NUMBER: 60/083,727
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/064,866
; PRIOR FILING DATE: 1997-11-05
; NUMBER OF SEQ ID NOS: 1558
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 181
; LENGTH: 13
; TYPE: RNA
; ORGANISM: Homo sapiens
; US-09-825-805-181
```

```

Query Match          100.0%; Score 7; DB 6; Length 13;
Best Local Similarity 57.1%; Pred. No. 6.1e+04;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 tgttga 7
    |::|||
Db 6 ugungga 12
```

```

RESULT 3
PCT-US01-30411-7
; Sequence 7, Application PC/TUS0130411
; GENERAL INFORMATION:
; APPLICANT: Genaisance Pharmaceuticals, Inc.
; APPLICANT: Choi, Julie Y
; APPLICANT: Kilem, Stefanie E
; APPLICANT: Koshy, Beena
; APPLICANT: Parks, Katie E
; APPLICANT: Stephens, J. Claiborne
; TITLE OF INVENTION: HAPLOTYPES OF THE TCF2 GENE
; FILE REFERENCE: TCF2_MMH-1439PCT
; CURRENT APPLICATION NUMBER: PCT/US01/30411
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/235,710
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US01-30411-7
```

```

Query Match          100.0%; Score 7; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 6.1e+04;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 tgttga 7
    |||||
Db 9 tgttga 15
```

```

RESULT 4
PCT-US01-30411-17/c
; Sequence 17, Application PC/TUS0130411
; GENERAL INFORMATION:
; APPLICANT: Genaisance Pharmaceuticals, Inc.
; APPLICANT: Choi, Julie Y
```

```

; APPLICANT: Kilem, Stefanie E
; APPLICANT: Koshy, Beena
; APPLICANT: Parks, Katie E
; APPLICANT: Stephens, J. Claiborne
; TITLE OF INVENTION: HAPLOTYPES OF THE TCF2 GENE
; FILE REFERENCE: TCF2_MMH-1439PCT
; CURRENT APPLICATION NUMBER: PCT/US01/30411
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/235,710
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US01-30411-17
```

```

Query Match          100.0%; Score 7; DB 1; Length 15;
Best Local Similarity 57.1%; Pred. No. 6.1e+04;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 tgttga 7
    |||||
Db 13 TGTGTGA 7
```

```

RESULT 5
US-09-825-805-182
; Sequence 182, Application US/09825805
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals, Inc.
; APPLICANT: Belgelman, Leo
; APPLICANT: Beaudry, Amber
; APPLICANT: Karpelsky, Alex
; APPLICANT: Adamic, Jasenka Matulic
; APPLICANT: Sweedler, Dave
; APPLICANT: Zinnen, Shawn
; TITLE OF INVENTION: Nucleotide triphosphate and their incorporation into oligonucleotides
; FILE REFERENCE: MBH800-831-F (400/009)
; CURRENT APPLICATION NUMBER: US/09/825,805
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 09/578,223
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 09/476,387
; PRIOR FILING DATE: 1999-12-30
; PRIOR APPLICATION NUMBER: 09/474,432
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 09/301,511
; PRIOR FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: 09/186,675
; PRIOR FILING DATE: 1998-11-04
; PRIOR APPLICATION NUMBER: 60/083,727
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/064,866
; PRIOR FILING DATE: 1997-11-05
; NUMBER OF SEQ ID NOS: 1558
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 182
; LENGTH: 15
; TYPE: RNA
; ORGANISM: Homo sapiens
US-09-825-805-182
```

```

Query Match          100.0%; Score 7; DB 6; Length 15;
Best Local Similarity 57.1%; Pred. No. 6.1e+04;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 tgttga 7
    |::|||
Db 7 ugungga 13
```

RESULT 6
US-09-527-030C-105
; Sequence 105, Application US/09527030C
; GENERAL INFORMATION:
; APPLICANT: VAN DOORN, Leen-Jan et al.
; TITLE OF INVENTION: Detection and identification of Human Papillomavirus by PCR and t
; FILE REFERENCE: 3501-0101P
; CURRENT APPLICATION NUMBER: US/09/527,030C
; CURRENT FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 497
; SOFTWARE: PatentIn version 3.0.
; SEQ ID NO 105
; LENGTH: 16
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Type specific probe derived from the Human Papillomavirus (HPV)
US-09-527-030C-105

Query Match 100.0%; Score 7; DB 6; Length 16;
Best Local Similarity 100.0%; Pred. No. 6.1e+04;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7
:|||||
Db 6 tgttga 12

RESULT 7
US-09-527-030C-181
; Sequence 181, Application US/09527030C
; GENERAL INFORMATION:
; APPLICANT: VAN DOORN, Leen-Jan et al.
; TITLE OF INVENTION: Detection and identification of Human Papillomavirus by PCR and t
; FILE REFERENCE: 3501-0101P
; CURRENT APPLICATION NUMBER: US/09/527,030C
; CURRENT FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 497
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 181
; LENGTH: 17
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic probe derived from the Human Papillomavirus (HPV)
US-09-527-030C-181

Query Match 100.0%; Score 7; DB 6; Length 17;
Best Local Similarity 100.0%; Pred. No. 6.1e+04;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7
:|||||
Db 7 tgttga 13

RESULT 8
US-09-685-664B-1636
; Sequence 1636, Application US/09685664B
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals, Inc.
; APPLICANT: Payco, Pam
; APPLICANT: McSwigen, Jim
; APPLICANT: Stinchcomb, Dan
; APPLICANT: Escobedo, Jaime
; TITLE OF INVENTION: Method and Reagent for Treatment of Diseases or Conditions Relate
; TITLE OF INVENTION: Levels of Vascular Endothelial Growth Factor Receptor

; FILE REFERENCE: MBH00-876-K (400/021)
; CURRENT APPLICATION NUMBER: US/09/685,664B
; CURRENT FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: US 60/005,974
; PRIOR FILING DATE: 1995-10-26
; PRIOR APPLICATION NUMBER: US 08/584,040
; PRIOR FILING DATE: 1996-01-08
; PRIOR APPLICATION NUMBER: US 09/371,772
; PRIOR FILING DATE: 1999-08-10
; NUMBER OF SEQ ID NOS: 8231
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1636
; LENGTH: 17
; TYPE: RNA
; ORGANISM: Homo sapiens
US-09-685-664B-1636

Query Match 100.0%; Score 7; DB 6; Length 17;
Best Local Similarity 57.1%; Pred. No. 6.1e+04;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7
:|||||
Db 10 uguuga 16

RESULT 9
US-09-685-664B-1637
; Sequence 1637, Application US/09685664B
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals, Inc.
; APPLICANT: Payco, Pam
; APPLICANT: McSwigen, Jim
; APPLICANT: Stinchcomb, Dan
; APPLICANT: Escobedo, Jaime
; TITLE OF INVENTION: Method and Reagent for Treatment of Diseases or Conditions Rel
; FILE REFERENCE: MBH00-876-K (400/021)
; CURRENT APPLICATION NUMBER: US/09/685,664B
; CURRENT FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: US 60/005,974
; PRIOR FILING DATE: 1995-10-26
; PRIOR APPLICATION NUMBER: US 08/584,040
; PRIOR FILING DATE: 1996-01-08
; PRIOR APPLICATION NUMBER: US 09/371,772
; PRIOR FILING DATE: 1999-08-10
; NUMBER OF SEQ ID NOS: 8231
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1637
; LENGTH: 17
; TYPE: RNA
; ORGANISM: Homo sapiens
US-09-685-664B-1637

Query Match 100.0%; Score 7; DB 6; Length 17;
Best Local Similarity 57.1%; Pred. No. 6.1e+04;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7
:|||||
Db 6 uguuga 12

RESULT 10
US-09-685-664B-3484
; Sequence 3484, Application US/09685664B
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals, Inc.
; APPLICANT: Payco, Pam
; APPLICANT: McSwigen, Jim
; APPLICANT: Stinchcomb, Dan

APPLICANT: Escobedo, Jaime
TITLE OF INVENTION: Method and Reagent for Treatment of Diseases or Conditions Related to Vascular Endothelial Growth Factor Receptor
FILE REFERENCE: MBH00-876-K (400/021)
CURRENT APPLICATION NUMBER: US/09/685,664B
CURRENT FILING DATE: 2000-10-10
PRIOR APPLICATION NUMBER: US 60/005,974
PRIOR FILING DATE: 1995-10-26
PRIOR APPLICATION NUMBER: US 08/584,040
PRIOR FILING DATE: 1996-01-08
PRIOR APPLICATION NUMBER: US 09/371,772
PRIOR FILING DATE: 1999-08-10
NUMBER OF SEQ ID NOS: 8231
SOFTWARE: PatentIn version 3.0
SEQ ID NO 3484
LENGTH: 17
TYPE: RNA
ORGANISM: Mus musculus
US-09-685-664B-3484

Query Match 100.0%; Score 7; DB 6; Length 17;
Best Local Similarity 57.1%; Pred. No. 6.1e+04;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 tqtgga 7
:|||||
DB 11 uguuga 17

RESULT 11
US-09-685-664B-3485
Sequence 3485, Application US/09685664B
GENERAL INFORMATION:
APPLICANT: Ribozyne Pharmaceuticals, Inc.
APPLICANT: Pavco, Pam
APPLICANT: McSwigen, Jim
APPLICANT: Stinchcomb, Dan
APPLICANT: Escobedo, Jaime
TITLE OF INVENTION: Method and Reagent for Treatment of Diseases or Conditions Related to Vascular Endothelial Growth Factor Receptor
FILE REFERENCE: MBH00-876-K (400/021)
CURRENT APPLICATION NUMBER: US/09/685,664B
CURRENT FILING DATE: 2000-10-10
PRIOR APPLICATION NUMBER: US 60/005,974
PRIOR FILING DATE: 1995-10-26
PRIOR APPLICATION NUMBER: US 08/584,040
PRIOR FILING DATE: 1996-01-08
PRIOR APPLICATION NUMBER: US 09/371,772
PRIOR FILING DATE: 1999-08-10
NUMBER OF SEQ ID NOS: 8231
SOFTWARE: PatentIn version 3.0
SEQ ID NO 3485
LENGTH: 17
TYPE: RNA
ORGANISM: Mus musculus
US-09-685-664B-3485

Query Match 100.0%; Score 7; DB 6; Length 17;
Best Local Similarity 57.1%; Pred. No. 6.1e+04;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 tqtgga 7
:|||||
DB 6 uguuga 12

RESULT 12
US-09-685-664B-3486
Sequence 3486, Application US/09685664B
GENERAL INFORMATION:
APPLICANT: Ribozyne Pharmaceuticals, Inc.

APPLICANT: Pavco, Pam
APPLICANT: McSwigen, Jim
APPLICANT: Stinchcomb, Dan
APPLICANT: Escobedo, Jaime
TITLE OF INVENTION: Method and Reagent for Treatment of Diseases or Conditions Related to Vascular Endothelial Growth Factor Receptor
FILE REFERENCE: MBH00-876-K (400/021)
CURRENT APPLICATION NUMBER: US/09/685,664B
CURRENT FILING DATE: 2000-10-10
PRIOR APPLICATION NUMBER: US 60/005,974
PRIOR FILING DATE: 1995-10-26
PRIOR APPLICATION NUMBER: US 08/584,040
PRIOR FILING DATE: 1996-01-08
PRIOR APPLICATION NUMBER: US 09/371,772
PRIOR FILING DATE: 1999-08-10
NUMBER OF SEQ ID NOS: 8231
SOFTWARE: PatentIn version 3.0
SEQ ID NO 3486
LENGTH: 17
TYPE: RNA
ORGANISM: Mus musculus
US-09-685-664B-3486

Query Match 100.0%; Score 7; DB 6; Length 17;
Best Local Similarity 57.1%; Pred. No. 6.1e+04;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 tqtgga 7
:|||||
DB 11 uguuga 7

RESULT 13
US-09-685-664B-3493/C
Sequence 3493, Application US/09685664B
GENERAL INFORMATION:
APPLICANT: Ribozyne Pharmaceuticals, Inc.
APPLICANT: Pavco, Pam
APPLICANT: McSwigen, Jim
APPLICANT: Stinchcomb, Dan
APPLICANT: Escobedo, Jaime
TITLE OF INVENTION: Method and Reagent for Treatment of Diseases or Conditions Related to Vascular Endothelial Growth Factor Receptor
FILE REFERENCE: MBH00-876-K (400/021)
CURRENT APPLICATION NUMBER: US/09/685,664B
CURRENT FILING DATE: 2000-10-10
PRIOR APPLICATION NUMBER: US 60/005,974
PRIOR FILING DATE: 1995-10-26
PRIOR APPLICATION NUMBER: US 08/584,040
PRIOR FILING DATE: 1996-01-08
PRIOR APPLICATION NUMBER: US 09/371,772
PRIOR FILING DATE: 1999-08-10
NUMBER OF SEQ ID NOS: 8231
SOFTWARE: PatentIn version 3.0
SEQ ID NO 3493
LENGTH: 17
TYPE: RNA
ORGANISM: Mus musculus
US-09-685-664B-3493

Query Match 100.0%; Score 7; DB 6; Length 17;
Best Local Similarity 100.0%; Pred. No. 6.1e+04;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tqtgga 7
|||||||
DB 14 tqtgga 8

RESULT 14
US-09-740-332-1100/C

```

; Sequence 1100, Application US/09740332
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals Inc.
; TITLE OF INVENTION: Enzymatic Nucleic Acid Treatment of Diseases or Conditions Relate
; TITLE OF INVENTION: Hepatitis C Virus Infection
; FILE REFERENCE: RPI 400/003
; CURRENT APPLICATION NUMBER: US/09/740.332
; CURRENT FILING DATE: 2001-03-26
; NUMBER OF SEQ ID NOS: 9704
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1100
; LENGTH: 17
; TYPE: RNA
; ORGANISM: artificial sequence
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION:
; OTHER INFORMATION: oligonucleotide substrate
US-09-740-332-1100

```

```

Query Match          100.0%; Score 7; DB 6; Length 17;
Best Local Similarity 100.0%; Pred. No. 6.1e+04;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 tgttga 7
        |||||
DB      16 TGTGGA 10

```

```

RESULT 15
US-09-740-332-1101/c
; Sequence 1101, Application US/09740332
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals Inc.
; TITLE OF INVENTION: Enzymatic Nucleic Acid Treatment of Diseases or Conditions Relate
; TITLE OF INVENTION: Hepatitis C Virus Infection
; FILE REFERENCE: RPI 400/003
; CURRENT APPLICATION NUMBER: US/09/740.332
; CURRENT FILING DATE: 2001-03-26
; NUMBER OF SEQ ID NOS: 9704
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1101
; LENGTH: 17
; TYPE: RNA
; ORGANISM: artificial sequence
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION:
; OTHER INFORMATION: oligonucleotide substrate
US-09-740-332-1101

```

```

Query Match          100.0%; Score 7; DB 6; Length 17;
Best Local Similarity 100.0%; Pred. No. 6.1e+04;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 tgttga 7
        |||||
DB      11 TGTGGA 5

```

Search completed: December 15, 2001, 05:58:20
 Job time: 12263 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 01:11:50 ; Search time 2725.73 Seconds
(without alignments)
8939.379 Million cell updates/sec

Title: US-09-380-826A-1

Perfect score: 1477
Sequence: 1 gatcagcgcacgaactaac.....ccgtaatcgatccctgcag 1477

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vl:*
30: em_hggo_hum:*
31: em_hggo_ity:*
32: em_hggo_rtd:*
33: em_htg_hum:*
34: em_htg_inv:*
35: em_htg_rtd:*
36: em_htg_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

8

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1431.4	96.9	1481	1	LFU60594	U60594 Leptospira
2	1388.6	94.0	1450	1	LEA19243	Y19243 Leptospira
3	1295.6	87.7	1508	1	LIRRN165	X17543 Leptospira
4	1270	86.0	1374	1	LII6SRDNY	221634 L.inadai ge
5	1263.8	85.6	1483	1	MA4516S	U12676 Leptospira
6	1263.8	85.6	1489	1	MA5216S	U12677 Leptospira
7	1262.2	85.5	1513	1	LBUI12670	U12670 Leptospira
8	1261	85.4	1489	1	LBUI12669	U12669 Leptospira
9	1260.8	85.4	1516	1	LNUI12671	U12671 Leptospira
10	1254.6	84.9	1425	1	WP4616S	U12630 L.borgpeter
11	1251.8	84.8	1486	1	LB16SRDNP	U12673 Leptospira
12	1246.6	84.4	1419	1	LK16SRDNP	U12628 L.kirschner
13	1239.8	83.9	1494	1	LSU12672	U12672 Leptospira
14	1238.2	83.8	1418	1	LM16SRDNP	U12648 L.meyerl ge
15	1232.8	83.5	1396	1	LII16SRRN	U12817 L.interroga
16	1230	83.3	1415	1	LS16SRDNP	U12649 L.santarosa
17	1225.4	83.0	1410	1	LM16SRDNP	U12637 L.wellll ge
18	1216	82.3	1424	1	LM16SRDNP	U12635 L.noguchii
19	1063.4	72.0	1523	1	LEPRR16SB	M88721 Leptospira
20	1031.8	65.9	1433	1	LB16SRDNP	Z98590 Leptospira
21	1027.6	66.6	1417	1	LM16SRDNP	Z21638 L.wolbachii
22	1022.2	69.2	1439	1	LB16SRDNP	Z98588 Leptospira
23	1003.6	67.9	1451	1	AA16SRDNP	M71241 Leptospira
24	1003.4	67.9	1321	1	LEPRR16SA	AF167353 Leptospir
25	1003.2	67.9	1366	1	AF167353	Z98587 Leptospira
26	999.8	67.7	1442	1	LB16SRDNP	Z21631 C.canela ge
27	995.8	67.4	1410	1	CC16SRDNP	Z12821 L.biflexa 1
28	993	67.2	1398	1	LB16SRDNP	Z26699 L.biflexa 1
29	991.8	67.1	1394	1	LB16SRDNP	Z98592 Leptospira
30	991.8	67.1	1394	1	LB16SRDNP	M88719 Leptospira
31	991.6	67.1	1379	1	LEPRR16SC	Z21633 J.Jequitilla
32	991	67.1	1379	1	LEPRR16SC	Z98593 Leptospira
33	985.8	66.7	1427	1	LB16SRDNP	Z98589 Leptospira
34	977.8	66.2	1466	1	LB16SRDNP	Z21632 L.illini ge
35	972	65.8	1445	1	LI16SRDNP	Z98591 Leptospira
36	964.4	65.3	1413	1	LB16SRDNP	AJ090479 unculture
37	915.8	62.0	1543	1	UEA19479	Z26770 L.biflexa g
38	884	59.9	1385	1	LB16SRDNP	U81652 Unidentifie
39	883.2	59.8	1475	1	UE081652	X77216 P.acidigall
40	874.4	59.2	1494	1	PA16SRRN	X75272 Clostridium
41	866.6	58.7	1494	1	CG16SRRN	Y17712 Malonomonas
42	866.6	58.7	1525	1	MR17712	X72869 C.thermopol
43	864.2	58.5	1503	1	CT16SRNA	AB039334 Brevibact
44	857.8	58.1	1532	1	AB039334	Z21636 L.parva gen
45	856.6	58.0	1379	1	LP16SRDNP	

ALIGNMENTS

RESULT	1	LOCUS	LFU60594	1481 bp	DNA	BCI	10-SEP-1998
DEFINITION		Leptospira fainei 16S ribosomal RNA gene, partial sequence.					
ACCESSION		U60594					
VERSION		U60594.1					
KEYWORDS		GI:1408219					
SOURCE		Leptospira fainei.					
ORGANISM		Bacteria: Spirochaetales: Leptospiraceae: Leptospira.					
REFERENCE		1 (bases 1 to 1481)					
AUTHORS		Perolat,P., Chappell,R.J., Adler,B., Baranton,G., Bulach,D.M., Billinghurst,M.L., Letocart,M., Merlen,F. and Serrano,M.S.					
TITLE		Leptospira fainei sp. nov., isolated from pigs in Australia					
JOURNAL		Int. J. Syst. Bacteriol. 48 Pt 3, 851-858 (1998)					
MEDLINE		98404550					
REFERENCE		2 (bases 1 to 1481)					
AUTHORS		Adler,B., Chappell,R.J., Baranton,G., Bulach,D.M., Billinghurst,M.L., Letocart,M., Merlen,F., Serrano,M.S. and Perolat,P.					

TITLE Direct Submission
JOURNAL Submitted (12-JUN-1996) Microbiology, Monash University, Wellington
RD, Clayton, VIC 3168, Australia
FEATURES Location/Qualifiers
source 1.1481
/organism="Leptospira fainei"
/strain="Hurstbridge"
/db_xref="taxon:48782"
rRNA <1..>1481
/product="16S ribosomal RNA"
BASE COUNT 391 a 335 c 439 g 314 t 2 others
ORIGIN

Query Match 96.9%; Score 1431.4; DB 1; Length 1481;
Best Local Similarity 99.7%; Pred. No. 1,4e-14;
Matches 1476; Conservative 0; Mismatches 1; Indels 4; Gaps 4;

QY 1 gatcatgctcaagaactcaagctgagcgcgctcttaacacatgaagtcgaagcgagtag 60
Db 1 GATCATGCTCAGAACTAACCTGGCGCGCTTTAAACTGCAAGTCAGCGGGGTAG 60

QY 61 caatactagcgcgcaagcggttgtaacacgt-gtcaatctccctccagagtcggagata 119
Db 61 CAATACCTAGCGCGCAAGCGGTGAGTAACAGTGTACTCTCCGAGCTCGGAGTA 120

QY 120 acttcgaaagaaactaatactacagatagctctgtgtgatacaagaattgata 179
Db 121 ACTTCCGAAAGAAAGCTAATACCGGATAGTCTGTTGGATCAACAGATTGTAGTA 180

QY 180 aagaattatgtcttgagatagacccgcgcgcgaatagctagtgtgtgaagtaagtgtc 239
Db 181 AAGATTATTGCTTGAGATGAGCCCGCGCGCGGATAGCTGTGAGGATATAGGCTC 240

QY 240 accaagcgagagatcagtagaccgctgagagaggtgtccgcgcaaatgagacgaaga 299
Db 241 ACCAAGCGAGAGATGAGTGGAGCGCTGAGAGGAGTCCGCGCAATGAGAACTGAGACA 300

QY 300 cgcgcatacaccctcagaggaagcgacagtlagaatctgtctcaatgagggaaacccctga 359
Db 301 CGGTCATACCTCTACCGGAGGACGACAGTTAAGATCTTGCTCAATGGGGAAACCCCTGA 360

QY 360 agcagcgagcgcgctgaaagaaagaggtctcgagatgtgaagttcaatgagcaag-gaa 418
Db 361 AGCAGCGACGCGCGGTGAAACGAAGAGCTTTCGATTGTAAAGTTCAATGCGCAGRGA 420

QY 419 aaataaagagaagaatgtatgtatgtaactgcttaagaacacggcgcaaatcagtcgaaga 478
Db 421 AAATAAGAGCAATGTGATGATGATGACTGCTTAAGACACCGGCTAATCTAGTCCACGA 480

QY 479 gccgcgtaatacgtatagtgcaagcgctgtctcgaaatcattggcgctgaaggtgcgta 538
Db 481 GCCGCGTAATACGTAATGTCGAACGCTTTCGGAATCATTTGGGCGCTAAAGGGTGCCTA 540

QY 539 ggcgagattgtcaagtcaagtggtgaanaactgcggtcgaaccggtggtcgtgaacttgaac 598
Db 541 GCGGAGATTGTAAGTCAAGTGTGAATACTGGCGGTCAACCGTGCTGCACTTGAAC 600

QY 599 tacaagtcgagatgttgagagagcgaaatggaatccaggtgtaagcggtgaatgcgtag 658
Db 601 TACAAGTGTGAGTTGGAGAGGCAATGGAATTCAGAGTGTAGCGGTGAATATCGTAG 660

QY 659 atatctgagagaacacagtgcgagagcgactgtcgctcaaaaactgaacgtgaagca 718
Db 661 ATATCTGAGAGAACACGAGTGGCGAGCGACTTCTGCTCAAAAACGACGCTGAGGCA 720

QY 719 cgaagcgctgggttagtaagcgatagatacccggtgaatccagacccctaaagttctc 778
Db 721 CGAAGCGCTGGGTATGAACGGGATTGATACCCCGGTATCCAGCCCTTAACGTTGTC 780

QY 779 taccagttgtgggggtttaaaccctcagtaacgaacctaagcaatgaatgaagcgcgt 838
Db 781 TACCAAGTTGTGGGGTTTAACCTCAGTAAGCAACTAAGCAATTAAGTAGACCGCT 840

QY 839 gggagactatgctcgcaagtgaaactcaaaaggaattgacgggggtccgcacaagcggtg 898
Db 841 GGGAGACTATGCTCGCAAGAGTAGAACTCAAAAGGATTTGAGGGGGTCCGCACAAGCGGTG 900

QY 899 gagcagtggtttaatctgatatgatacccaaaaacccctaccctggtgcttgacatgactg 938
Db 901 GAGCATGTGCTTAAATTCAGTAGATACCCCAAAAACCTCACCTGGGCTTGAATGAGATCTG 960

QY 959 aatcagtagagatatatgagccttcggcgagatcgaagtgtagatggttgatgca 1018
Db 961 AATCATGTAGAGATATATGAGCTTCGGCGAGATTCAACAGTGTGATGATGCTGCA 1020

QY 1019 gctcgtgctgtaagatgttggttaagtcgccgaagcgcaacccctacgtatgtg 1078
Db 1021 GCTCGTGTGCGAGATGTTGGTTAAAGTCCCGCAGACGCGCAACCCCTATGCTATGTTG 1080

QY 1079 ctacc-ttaagttgagcactgtgtaagaactgcggtgtaacaaacccggaaggaagcgagga 1137
Db 1081 CTACCATTAAAGTTGGGCACTGCTAGCAAAACTGCCGTGACAAACCGGAGGAGGCGGGGA 1140

QY 1138 tgaagtaaatccctcatalgagccttatgtccagggcacaacacgctgctacaatgagcgata 1197
Db 1141 TGAGGTCAAAATCTCATAGGCTTTATGTGCCAGGSCACACAGTGTCAATGAGCCGATA 1200

QY 1198 cagagagtcgcgaactcgcaagagaggaactaactcctaaagtcggtccaggtcgagtt 1257
Db 1201 CAGAGGTCGCGCAACTCGCAAGAGGAGGACTAATCTTAAAGTGGTCCCGAGTTCGAGTT 1260

QY 1258 ggggtctgcaactcgaccccaatgaaatcgagatcgctagatagtcggatagaatgacg 1317
Db 1261 GGGGTCTGCAACTCGACCCCAATGAAAGTCGAATCGTAGTAATCGGATAGCATGCGC 1320

QY 1318 cgtgtaatacgtcccggaactgttacacacgcgcgtcaacacacactgagttgggagca 1377
Db 1321 CGGTGAATACGTTCCCGAGCTTGTACACACGCCCGTCAACACCGATGGGAGCA 1380

QY 1378 cccgaagtgtctctgttaacgtaagagagagactaactaagttgaactcgtgaaggg 1437
Db 1381 CCCGAAGTGTCTTGTGTTAAACGTAAGGAGACAGACTAAGGTGAAACGCTGAAGGG 1440

QY 1438 ggtgaagtcgtaacaagta-ccgtaaatgatacttcctgag 1477
Db 1441 GGTGAAGTCGTAAACAAGGTARCCGTAATCGATTCTGACG 1481

RESULT 2
LFA19243
LOCUS LFA19243 1450 bp DNA BCT 08-NOV-1999
DEFINITION Leptospira fainei partial 16S rRNA gene, SSI 5402-98.
ACCESSION Y19243
VERSION Y19243.1 GI:6318184
KEYWORDS 16S ribosomal RNA; 16S rRNA gene.
SOURCE
ORGANISM Leptospira fainei.
REFERENCE 1 (bases 1 to 1450)
AUTHORS Petersen, A.M., Krogfelt, K.A., Perolat, P., Boye, K. and Schlichting, P.
TITLE Leptospira fainei serovar Hurstbridge isolated from two Danish patients with Weil's syndrome
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1450)
AUTHORS Krogfelt, K.A.
TITLE Direct Submission
JOURNAL Submitted (05-NOV-1999) K.A. Krogfelt, Statens Serum Institut, Dept. of Gastrointestinal Infections, Artillerivej 5, 2300 Copenhagen, DENMARK
FEATURES
source 1.1450
/organism="Leptospira fainei"
/strain="SSI 5402-98"
/db_xref="taxon:48782"

rRNA
 /gene="16S rRNA"
 /product="16S ribosomal RNA"
 gene
 1.1450
 /gene="16S rRNA"
 BASE COUNT 383 a 326 c 437 g 304 t
 ORIGIN

Query Match 94.0%; Score 1388.6; DB 1; Length 1450;
 Best Local Similarity 99.2%; Pred. No. 5.9e-14;
 Matches 1427; Conservative 0; Mismatches 9; Indels 3; Gaps 3;

Oy 24 ggcggcgccgttctaacaatgcgaatgcgaacggggtgagcaataccttaagcggcgaacgggtg 83
 Db 1 ggcggcgccgttctaacaatgcgaatgcgaacggggtgagcaataccttaagcggcgaacgggtg 60
 Oy 84 agtaacacgtt-gtaatactctccagagcttgagataactctcgaagaaggaagtaata 142
 Db 61 agtaacacgttggtatcttctcctccagagcttgagataactctcgaagaaggaagtaata 120
 Oy 143 ccggaatagctctgttgatcacagaattgataagtaagaattatcttgaggatgag 202
 Db 121 ccggaatagctctgttgatcacagaattgataagtaagaattatcttgaggatgag 180
 Oy 203 ccggaatagctctgttgatcacagaattgataagtaagaattatcttgaggatgag 262
 Db 181 ccggaatagctctgttgatcacagaattgataagtaagaattatcttgaggatgag 240
 Oy 263 ggcctgagaaggtgtccggaacacatggaactgagacaagtcctactactcgaagagc 322
 Db 241 ggcctgagaaggtgtccggaacacatggaactgagacaagtcctactactcgaagagc 300
 Oy 323 agcagtttaagaatcttgcctcaatgagggaaccccggaacgagcggcggtgaaagaa 382
 Db 301 agcagtttaagaatcttgcctcaatgagggaaccccggaacgagcggcggtgaaagaa 360
 Oy 383 gaaggtcttcgagatgtaaatgtaacttaagcagg-aaataagcagcaatgtgatgac 441
 Db 361 gaaggtcttcgagatgtaaatgtaacttaagcaggaaataagcagcaatgtgatgac 420
 Oy 442 gtacctgccttaagaacacggcttaactacgtgcagcagcggcgttaactacgtatgta 501
 Db 421 gtacctgccttaagaacacggcttaactacgtgcagcagcggcgttaactacgtatgta 480
 Oy 502 agcgtgttccgaatcatcttgagcgtgaaaggtgctgagcggatctgttaagccaagttg 561
 Db 481 agcgtgttccgaatcatcttgagcgtgaaaggtgctgagcggatctgttaagccaagttg 540
 Oy 562 aaaaactgcggagctcaaacccgtgacctgaaactcaactcaagctcggagcttgaggagag 621
 Db 541 aaaaactgcggagctcaaacccgtgacctgaaactcaactcaagctcggagcttgaggagag 600
 Oy 622 gcaagtggaatctccagtgtagcgggtgaaatgtagatatacttgaggaacacccagtgagc 681
 Db 601 gcaagtggaatctccagtgtagcgggtgaaatgtagatatacttgaggaacacccagtgagc 660
 Oy 682 gaagcgagacttgctggttcaaaactgagcgtgagcagcaaaagcgtgggtagtaaacggg 741
 Db 661 gaagcgagacttgctggttcaaaactgagcgtgagcagcaaaagcgtgggtagtaaacggg 720
 Oy 742 atagataccccggtaactcaacgccccttaaacgttctcaccagcttgcttgagggtttaac 801
 Db 721 atagataccccggtaactcaacgccccttaaacgttctcaccagcttgcttgagggtttaac 780
 Oy 802 cctcagtaacgaacctaacagtaataagtagaccgcctgaggactatgctcgcgaagaatga 861
 Db 781 cctcagtaacgaacctaacagtaataagtagaccgcctgaggactatgctcgcgaagaatga 840
 Oy 862 aactcaagaagatgacgggggtccgcacaaagcgtgtgagcgtgtggtttaaattcatga 921
 Db 841 aactcaagaagatgacgggggtccgcacaaagcgtgtgagcgtgtggtttaaattcatga 900

Oy 922 taocccaaaacctcaacctgggcttgatagatgatactgtaactgtagagatatatgacc 981
 Db 901 taocccaaaacctcaacctgggcttgatagatgatactgtaactgtagagatatatgacc 960
 Oy 982 ttccggcagatctcaacaggtgtcgtgagtggttgcgtgaactcgtgtcgtgagatggtggt 1041
 Db 961 ttccggcagatctcaacaggtgtcgtgagtggttgcgtgaactcgtgtcgtgagatggtggt 1020
 Oy 1042 taagtcgccgaacggcggaacccctatcgtatgtgtgacc-ttaagttggcactggt 1100
 Db 1021 taagtcgccgaacggcggaacccctatcgtatgtgtgacc-ttaagttggcactggt 1080
 Oy 1101 acgaacatcgctgagcaaacccggaagggcgggagtagtaactcaatccatcctgacct 1160
 Db 1081 acgaacatcgctgagcaaacccggaagggcgggagtagtaactcaatccatcctgacct 1140
 Oy 1161 tatgtccaggcgcacacacgtgtctaacatgagcagatacagaaggttcgccaactcgaaga 1220
 Db 1141 tatgtccaggcgcacacacgtgtctaacatgagcagatacagaaggttcgccaactcgaaga 1200
 Oy 1221 gggagtaactctcctaagaatggtgccagcttgagatggtgtcgtcaactcgaccatg 1280
 Db 1201 gggagtaactctcctaagaatggtgccagcttgagatggtgtcgtcaactcgaccatg 1260
 Oy 1281 aagtcggaatcgctagtaatcgagatcgagatgcgagtgagtaactcgtcccgacct 1340
 Db 1261 aagtcggaatcgctagtaatcgagatcgagatgcgagtgagtaactcgtcccgacct 1320
 Oy 1341 gtacacacccgcgttcacacacacgtgagtgaggagcaccggaagtggtcttgtaaccg 1400
 Db 1321 gtacacacccgcgttcacacacacgtgagtgaggagcaccggaagtggtcttgtaaccg 1380
 Oy 1401 taagggagacagctactaagtggaactcgttaaggggtggaagtggttaacgaagtaac 1459
 Db 1381 taagggagacagctactaagtggaactcgttaaggggtggaagtggttaacgaagtaac 1439

RESULT 3
 LIRN16S
 LOCUS LIRN16S 1508 bp DNA BCT 04-SEP-1991
 DEFINITION Leptospira interrogans 16S ribosomal RNA gene.
 ACCESSION X17547
 VERSION X17547.1 GI:44008
 KEYWORDS 16S ribosomal RNA; ribosomal RNA.
 SOURCE Leptospira interrogans.
 ORGANISM Leptospira interrogans.
 Bacteria; Spirochaetales; Leptospiraceae; Leptospira.
 REFERENCE
 1 (bases 1 to 1508)
 AUTHORS Fukunaga, M.
 TITLE Direct Submission
 JOURNAL Submitted (22-NOV-1989) Fukunaga M., Faculty of Pharmacy and
 Pharmaceutical Science, Fukuyama University, 965 Higashimuracho,
 Fukuyamashi, Hiroshima 729-0, Japan
 2 (bases 1 to 1508)
 AUTHORS Fukunaga, M., Horie, I., Okuzako, N. and Mifuchi, I.
 TITLE Nucleotide sequence of a 16S rRNA gene for Leptospira interrogans
 serovar canicola strain Moulton
 JOURNAL Nucleic Acids Res. 18 (2), 366 (1990)
 MEDLINE 90221824
 COMMENT Data kindly reviewed (23-FEB-1990) by Tsuchida M.
 FEATUSES Location/Qualifiers
 source 1. 1508
 /organism="Leptospira interrogans"
 /strain="serovar canicola strain Moulton"
 /db-xref="taxon:173"
 /note="16S ribosomal RNA"
 BASE COUNT 393 a 342 c 456 g 317 t
 ORIGIN

Query Match 87.7%; Score 1295.6; DB 1; Length 1508;
 Best Local Similarity 94.7%; Pred. No. 1.3e-12;

Matches	1384:	Conservative	0:	Mismatches	74:	Indels	4:	Gaps	4:
OY	1	gacatcgtcagaactaacgctcggcgcgctccttaacatgcaagtcgagcggtgag	60						
Db	15	GATCTGGGCTCAAGAACTAACGCTGGCGCGCGCTTAAACATGCAAGTCAAGCGGATAG	74						
OY	61	caatacttaacgagcagcggtgagtaacacgt-ggtaactcttcctccgcttgagata	119						
Db	75	CAATACCTACAGCGGGAAGGGGTAGTACACGCTGGTAACTCTTCTGTGAGCTGGGATA	134						
OY	120	acttcgcaaaaggaagctaaataccggatagtcctctgttgatacaagatttgata	179						
Db	135	ACTTTCGGAAGGGAACCTAATACCTGATGTCGCCAGAGATCAAGATTTTTCGGGTA	194						
OY	180	aagatttaattgtgtgagatgagagcccgccgataagctagttgtgtgagatagctc	239						
Db	195	AAGTATTAATGCTCGAGATGAGCCCGCTCCGATTAGCTAGTGGTGAAGTAAGGCTC	254						
OY	240	accaaagcagacatcgtgtaacggcctgagaggtgtccggccaacatggaactgagaca	299						
Db	255	ACCAAGCGAGATCGGTAGCGCGCTGAGAGGGTGTTCGGCCCAATGGAACTGAGACA	314						
OY	300	cgtgcatactcctaagggagagcagcaatgaatacttctgctcaatggggaaacctga	359						
Db	315	CGGTCCATACCTCTACCGGAGCGCAGATTAAATCTTCTCAATGGGGGAA-CCTGA	373						
OY	360	agcagcagacgcccgtgaaacgaagaagttcgcgattgtaaatgtaattcatagaca-g	418						
Db	374	AGCAGCGACGCGCGGTGAACATGAAGTCTTCGGATTGTAAATTTCAAGTAAAGGAA	433						
OY	419	aaataagcagaatgtgatgtgtgtaactgcctaaagcaccggttaactacgtgccaga	478						
Db	434	AAATAAGAGCAATGTGATGTGATGCTGCTTAAGACACCGGCTAATACGTCCAGCA	493						
OY	479	gcgcggttaaacgtatggtgtcaagcgtgtgttcggaatcatctggcgtlaaaggtg	538						
Db	494	GCCCGGTAATATCATGTGTGCAAGCGCTTGTTCGAAATCTTGGCGCTAAAGGGTG	553						
OY	539	ggcgattgtgaagtcaggtgtgaacactgcggtcaaccgctgcctgaacttgaac	598						
Db	554	GGCGGACATGTAAGTCAAGTGTGAACCTGGGGCTCAACTCGCAGCTGCACTTGAAC	613						
OY	599	tacaagtcgtgagtttggagagcgaatgtgaatcagaagtgtgtagcggatgagctag	658						
Db	614	TATGTGCTCGAGTTTGGAGAGCGCAATGTGATTCAGGTGTGAGCGGTGAATCCG	673						
OY	659	atactcggagaaacacacggtgagaaagcgactgtcgtgcctaaacgaagcgtgagaa	718						
Db	674	ATATCTGGAGGAACCACTGGCGAGGCGACTTGTGCTTAACACTGACGCTGAGGCA	733						
OY	719	cgaagcgtggtgtagtaaacgggattagatacccggttaatccaacgcccctaaacgt	778						
Db	734	CGAAAGCTGGGTAGTGAAGGATGATTAATACCCCGGTAATTCAGCCCTTAACGTTG	793						
OY	779	tacaagttgttgggggttttaacctcaagtaacgaacttaacgattagtagacgctc	838						
Db	794	TACAGATTGTTGGGGGTTTAAACCTCACTAAGCAACCTAACGGGTTAAGTAGACCG	853						
OY	839	ggggaactatgctcgaaagtgaaactcaaggaattgacgggggttcgcgcaaacggtg	898						
Db	854	GGGGGCTATGCTCGCAAGAGTGAACCTCAAGAACTTAACGGGGGTCGGCACAGCGGT	913						
OY	899	gagcactggtttaattcgaatgatacccaaaaactcaactcactggtgctgacatgata	958						
Db	914	GAGCATGTGTTTATTCGATGATACGGGAAACCTCACTAGGCTTGACATGAGTAGG	973						
OY	959	aatcatgtagagataatgagccttcggcgagattcaacagtgctgcgatgtgtcgctca	1018						
Db	974	AATCATGTAGAGATACATGAGCTTCGGGCGCTTCACAGGTCGTGATGTTGTCGCA	1033						
OY	1019	gctcgtgtcgtgagatgttgggttaagtcccgcaacgagcgcaacccctatgata	1078						
Db	1034	GCTCGTGTGATGATGTTGGGTTAAGTCCGCAACGAGCGCAACCTCACTTATGTTG	1093						

OY	1079	cta-cccttaagttgggcactggtacgaaactgcccgttgacaacacggaggaagcgggga	1137						
Db	1094	CAATCATTCAGTTGGGACTTCGTAGGAAGTCCCGGTGACAAACCGGAGAACGGGGGA	1153						
OY	1138	tgaagtcataatccctcatgagccttaatgctcgaagccacacagctgtacaaatggcgata	1197						
Db	1154	TGAGTCGAATTCCTCATGAGCGCTTATGTGTAGGGCAACACACGTGTACAAATGGCCG	1213						
OY	1198	cagaaggttcgcaacttcgcaagggagagctaatctctaagaacggttcccgatcgatt	1257						
Db	1214	CAAGGGTATCCCACTTCGACGAGGGGAGCTTAATCTCAAAATCCGGTCCAGTTCCGATT	1273						
OY	1258	ggggttcgaactcgaccccatgaagtcggaatcgctagtaatcgcggaatcaagatgcg	1317						
Db	1274	GGAGTTCGAACCTGCATTCATGATGAGTGGGAATCGTAAATCGCGATCAGCATGCCG	1333						
OY	1318	cgtgtaactgctcccgagcctgtgatacacacgcccgttaacacacactgtgtggagca	1377						
Db	1334	CGGTGAATACGTTCCCGACCTTGTACACACCGCCCTCACACCTGAGTGGGAGCA	1393						
OY	1378	cccgaatggtcttctgttgaacggtgaagagacagactactaaagtgaactcgtaaagg	1437						
Db	1394	CCCGAATGCTCTTTGCCAACCGCAAGGAAGACAGACTACTAAGTGAATCTCTGAAAGG	1453						
OY	1438	ggtgaagtcgtlaacaaggtacc	1459						
Db	1454	GGTGAAGTCGTAAACAGGTAGC	1475						
RESULT	4								
L116SRNY									
LOCUS	L116SRNY	1374 bp	DNA	BCT					02-DEC-1993
DEFINITION	L116SRNY	gene for 16S ribosomal RNA (partial).							
ACCESSION	221634								
VERSION	221634.1	GI:433581							
KEYWORDS	16S ribosomal RNA.								
SOURCE	Leptospira inadal.								
ORGANISM	Leptospira inadal								
REFERENCE	1 (bases 1 to 1374)								
AUTHORS	Hooker, J.V.								
TITLE	Phylogeny of Leptospiraceae and related Spirochaetes								
JOURNAL	Unpublished								
REFERENCE	2 (bases 1 to 1374)								
AUTHORS	Hooker, J.V.								
TITLE	Direct Submission								
JOURNAL	Submitted (09-FEB-1993) Hooker J.V., Public Health Laboratory								
	Service, Leptospira Reference Laboratory, Stonebow Road, Hereford,								
	Herefordshire, United Kingdom, HR1 2ER								
FEATURES									
source	1. 1374								
	/organism="Leptospira inadal"								
	/strain="LYME"								
RNA	<1..>1374								
	/product="16S ribosomal RNA"								
BASE COUNT	358 a 311 c 413 g 289 t								3 others
ORIGIN									
Query Match	86.0%;	Score 1270;	DB 1;	Length 1374;					
Best Local Similarity	96.5%;	Pred. No. 3.4e-12;							
Matches 1327:	Conservative	2;	Mismatches 43;	Indels 3;	Gaps 3;				
OY	16	ctaagcgtgagcgcgctcctaatacatgcaagtcagcggggtagcaatactagcgcg	75						
Db	1	STARCGTTGGCGCGCGCTTAAACATGCAAGTCAAGCGGGGTAGCAATACTAGCGCG	60						
OY	76	aacggtgagtaacacgt-ggtaactctctcgcgagtcctgggataacttcgaaagaa	134						
Db	61	AACGGGTAGTACACAGTGGTAATCTTCTCCGAGTCTGGGATTAATCTCCGAAGGAA	120						

OY	135	agcctaaaccggaatgctcgttttggaaccagaatttgaatgagtaaaattatctgctt	194
Db	121	AGCTAAATACCGGATAGTCTCTACTGATCTACAGAGATCTGTATAGTAAATTTATGCTTG	180
OY	195	gagaatgaccgcygcgcgaattagctatgtgtgaagttaatgtgtctcaagaagcygacatc	254
Db	181	GAGATGAGCCCGCGCGCGGATTAAGCTTGTTGCTGAGGTAAAGGCTCCACCAAGCAGCATC	240
OY	255	gtatgcgcgccttgaaggagtgctccggcccaaatggaaactgagaaacggttcaatctcta	314
Db	241	GGTAGCCGCGCTGAGAGGtGTCCGCCCAATGGAAGTGAACAGCGGTTCATCTCTTA	300
OY	315	cggaaggcagcagttlaagaaattctgtcctaaatgggggnaaaccttgaagcagcgaacgcgcg	374
Db	301	CGGAGGCACACAGTTAAAGATCTTGCTCAATGAGGGGAAACCTGAAAGCAGGACGCCGC	360
OY	375	tgaacgaagaagtgcttcggatgtgtaagttaattcaatgaca -ggaataataagcaaatg	433
Db	361	TGAACGAGAAGAGCTCTTGATTTGTAAGTTTAAATTAAGCAGGAAAAATTAAGCAGCATG	420
OY	434	tgaatgagtaacctgcctaaagacccggttaactagctgbcagcagccgcgtaatacgt	493
Db	421	TGATGATGATCTCTGCTTAAGACCGGCTTAACCTACGTACGTCCAGCAGCCGCGTAAATAGT	480
OY	494	atggtgcaagcgttgttcggaatcatltgggcgtlaaagggtgcgtatgagcgtttgtaagt	553
Db	481	ATGTGCAACAGTTGTTCGGAATCATTTGGGCGTAAAGGTCGCGTAGCGGATTTGTAAAT	540
OY	554	caggtgtgaaaaactgcygcgtctaacccggtgacctgaaactgaataaactcgtgagtc	613
Db	541	CAGGTGTGAANAACGCGGGCTCAACCCGtgGCGCTGACTTGAACCTACAABCTGTGAGTT	600
OY	614	tggagaagcgaatlggaattccacagcgttagcgcgtgaaatgtcgaatactcgtgaagaca	673
Db	601	TGGAGAGCGCAAGTGAATCCAGGtGTAGCGGTGAATAGCTAGATATCTGGAGGAACA	660
OY	674	ccaatggcgaagcgcgacttgctggtcccaaaactgacgcgtgagagccgaagacgttgggttag	733
Db	661	CCAGTGCAGAAAGGAGATTCTCTCACTCAAAACCTGACGCTGAAGGACGAAACGtGGTAG	720
OY	734	taaacgggaattagaataccccggttaattccacgcgccttaaacggttgcatacagtttgtggg	793
Db	721	TAAACGGGATTTAGTATACCCCTGTGAATCCAGCCGCTTAACGTTGTCTTACCACTTGTGGGG	780
OY	794	gttttaacctcagttaacgaacctaaaggaattaagttagacgcgcttgggactatgctgc	853
Db	781	GTTTTAACCTCTCAATTAACGAACCTTAACGGATTAAAGTAGACCCCTCGGGAGCTATGCTCGC	840
OY	854	aagagtbaaaactcaaaaggaattagaacggggttcgcgaacgaagcgttgagcatgtggttaa	913
Db	841	AAGAGTAAACTCAAAAGAAATTACCGGGGCTCCGACAAAGCCGTGTGACATATGTGGTTTAA	900
OY	914	ttcgaatgaatacccaaaaactcaacctcgttggtcttgacatlgatccttaacatgtagagata	973
Db	901	TTTCATGTATACGGCAAAAACCTCACTCGGGGCTTGAACATGATCTGAATCATGTAGAGATA	960
OY	974	tatgagccttcgggcagatctcaagtgctctgaatgtgttgctcaagctcgtgtcgtgaga	1033
Db	961	TATAGAGCTTTCGGCAGATTTCACAGGTGTGATGTGTTCTCTACGCTCGTGTGTGAGA	1020
OY	1034	tgtttggttlaagtcgccgaacgagcgaacccctatcgtatgttgcatacctaagtbtgg	1093
Db	1021	TGTTT -GGTTAAGTCCGCCAAGAGGCGCAACCTCACCTTATGTGGCATCATTTCAAGTTGGG	1079
OY	1094	caactgttaacgaacttcgcggttgacaacacgggaaggaagcgggagtatgactcaaatctcta	1153
Db	1080	CACTCGTAAAGAAATCGCGGTGACAAACCGAGAAAGCGGGGTGTAGACGTCAAGTCCTCA	1139
OY	1154	tggccttatgttccaagcgcacacagttgcttaaaatgycgcgtataagaagggttcgccaact	1213
Db	1140	TGGCTTTATGTCTTAGGGCAACACAGCTGCTCAATAGCGCGGTACAAAGGTAAGCAACT	1199
OY	1214	cgcaagagggaactaactcttaaaagtgcgtccagttcgaattggggtctgacaactctga	1273

QY	1274	ccccatgaagtcggaatcgcgtagtaatcgcggatcagaatgcgcggatgaaatcgtccc	1333
Db	1260	CTCCATGAAAGTCGGAATCGTAGTATCGCGGATCAGCATGCCCGGTGAATACGTTCC	1319
QY	1334	ggacctttacacaccgcgcgtcaacacaccttgatgggagacccccaagtgt	1388
Db	1320	GGACCTTTACACACCGCCCGTCACACCACTGAGTGGGGAGACCCGAAAGTGGT	1374
RESULT	5		
LOCUS	WA4516S	1483 bp	DNA BCF 18-FEB-1995
DEFINITION	Leptospira weilii Cellidoni (ATCC 43285) 16S rRNA gene, partial sequence.		
ACCESSION	U12676		
VERSION	U12676.1	GI:558937	
KEYWORDS			
SOURCE	Leptospira weilii.		
ORGANISM	Leptospira weilii		
REFERENCE	Bacteria; Spirochaetales; Leptospiraceae; Leptospira.		
AUTHORS	1 (bases 1 to 1480)		
TITLE	Ralph,D. and McClelland,M.		
JOURNAL	Phylogenetic evidence for horizontal transfer of an intervening sequence between species in a spirochete genus		
MEDLINE	J. Bacteriol. 176 (19), 5982-5987 (1994)		
REFERENCE	95014031		
AUTHORS	2 (bases 1 to 1483)		
TITLE	McClelland,M.		
JOURNAL	Direct Submission		
FEATURES	Submitted (25-JUL-1994) Michael McClelland, California Institute of Biological Research, 11099 North Torrey Pines Road, La Jolla, CA 92037, USA		
source	location/qualifiers		
	1..1483		
	/organism="Leptospira weilii"		
	/strain="serogroup Cellidoni, serovar cellidoni, strain		
	/cellidoni (ATCC 43285)"		
	/isolate="WA45"		
	/db_xref="taxon:28184"		
	<1..1483		
	/product="16S rRNA"		
BASE COUNT	382 a 331 c 452 g 315 t 3 others		
ORIGIN			
Query Match	85.6%;	Score 1263.8;	DB 1; Length 1483;
Best Local Similarity	94.1%;	Pred. No. 3.9e-12;	
Matches 1376;	Conservative	0; Mismatches	80; Indels 6; Gaps 6;
QY	1	gatactgctcgaagactaacgctgctgcgcgcgtcttaacaatgcgaagcgggtag	60
Db	10	GATCCTGGCTCAGAACTAACGCTGGCGGCGGTCTTAACAATGCAAGTCAAGCGGAGTAG	69
QY	61	caaatccttagcggcgaaecgggtgagtaacaagt-ggtatactctccgcgagctggagta	119
Db	70	CAATCTCAGCGCGCAACGGGTGAGTAACAGCTGGGTAACTCTCCGAGCTCGGATTA	129
QY	120	acttcgcaaaagaagctaataacacggatagtcctgttggatcacaagaattgata	179
Db	130	ACTTTCGCAAAAGGAGCTAATACCTGATGATGCCCGAAGAGTCAATATGATTTTCGGGTA	189
QY	180	aagattatctgttggagatgagccgcgcgcgcatltagctatgttggatgagtaatgctc	239
Db	190	AAGATTATTTGCTCGGAGCTGAGCCCGGCCCGATTAAGTAACTAGTTGCTAGGTAAAGCTC	249
QY	240	accagaagcagacatcgtgtagccggcgtctgaaaggtgtcggccacaatggagacga	299
Db	250	ACCAAGCGCAGCATGCGTAGCGCGCTGAGAGAGGTGTTCCGCCCAATGCAATGACACA	309
QY	300	cggtcatactctcctacggaagcagcagttaagaatcttgcctaattgggggaaacctga	359

```

||||| 310 CGGTCCATCT-CTACGGAGGCGACGATTAAGATCTTCTCATATGGGGAAACCTCTGA 368
||||| 360 agcagcgacgcgcgctgaaagaaaggcttcggatlttaagtaagtaagca-ggaa 418
||||| 369 AGCAGCAGCGCGCGTGAACGATGAGTGCCTTCGATTTGAAATGTAATGAACGAGGAA 428
||||| 419 aaataagcagaatgtagatgtagtacctgctaagaacacgggttaacttcgtccagca 478
||||| 429 AAATAACACAGGATGATGATGTGATCTTAAGCACCGGCTTAACCTAGTCCAGCA 488
||||| 479 gccgcggtlaaagcagtcagtcgaaagcgttcgtaacatltggcgctlaaaggctgcta 538
||||| 489 GCCCGGTAATACCTATGCTGCAAGCGCTTGGGAATCATTTGGCGCTGAAGGCTGCGTA 548
||||| 539 gccgagatttgaagtcaggtgctgaagaacgcggcctcaaccgctgacctgaagc 598
||||| 549 GCGGACATGTAATCAAGTGTGAACCTCGGGCTCAACTCGCAGCTGCATTGTAAC 608
||||| 599 taacaagtcgaagcttgggaagcgaatggaattccagcggtgagcggtgaagtcgtag 658
||||| 609 TATGCTGTGAGTGTGGAGAGGCAAGTGAATTCAGGCTGAGCGGTGAATGCGGTAG 668
||||| 659 atatcgtgaagaaacacccgctggcgaagcgacttgcgtgctcaaaactgacgctgaagca 718
||||| 669 ATATCTGAGAGAACCCAGTGCAGGCGACTTGGCTTAATAACTGACGCTGAGAGCA 728
||||| 719 cgaagcgctggtgtagtaaacgggattagatacccggtgaatccagcgccttaaaagttgct 778
||||| 729 CGAAAGCTGGGTGTGAGTGAACGGATTAATACCCGCTTAATCCAGCCCTTAACGCTGTCTC 788
||||| 779 taacagctgttcgaggtgtttaaaccctcagtaacgaacctaaacgtaagtagaccgct 838
||||| 789 TACCGATTGTTGGGGGTTTAACCTCAGTAACGAA-CTAACGGATTAAGTAAGACCGCT 847
||||| 839 gggagactatgctcgcaagatgaaactcaaaaggaattgacgggggtccgcacaaagcggt 898
||||| 848 GGGGACTATGCTCGCAAGAGTGAACCTCAAGGAATTGACGGGGGTNNNCAACAAACGGGTG 907
||||| 899 gaagcattggtttaattatgatgatatacccaaaaacatcaccctgagcttgaatgtagctg 958
||||| 908 GAGCATGTGTTTAATTTGATGATGATACGCGMAAACCTCACTAGCTTGACATGAGAGTGG 967
||||| 959 aatcatgtagaagataatagacccctcggcagatcacaggtgctgacatggtgtgctca 1018
||||| 968 AATTATGTAGATACATGAG-CTTCGGGCGCTTACAGAGGTGCTGATGTTGTCGCA 1026
||||| 1019 gctcgtgctgtagatggttgggttaagtcctccgcaacgagcgcaaccctatcgtatgtt 1078
||||| 1027 GCTCGTGTGAGATGTTGGGTAAAGTCCCAACGACGCGCAACCCCTCACCTTATGTTG 1086
||||| 1079 cta-ccttaagctggcagctgtaacgaactlccggtgacaaacgcggaggaagcgagga 1137
||||| 1087 CCATCATTTAGTTGGGCACTGCTAAGGAACCTCGGTGCAAAACGGAGAGGCGGGGA 1146
||||| 1138 tgaagtcataatcccaatgagccttattgctcagggcacaacacgctgctacaatggcgata 1197
||||| 1147 TCACGCTCAATCTCATGCTCTTTATGCTTAGGCAACACAGTGTCTACAAAGCGGGTA 1206
||||| 1198 cagagagctgcacaactcgaagaggaagtaatctcctaagaatcggtlcccaatcgagtt 1257
||||| 1207 CAAAGGCTAGCCAACTCGGAGGGGAGCTAATCTCAAAAACCGGCTCCCACTTGGGATT 1266
||||| 1258 ggggtctgaactcgaccccaatgaagtcggaatcgctagtaatcgcgatcagcatgccc 1317
||||| 1267 GGAGTCTGCAACTCGACTCATGAAGTCCGAATCCCTAATAATCGCGATGACATGCCG 1326
||||| 1318 cggatgaatcgctcccgagctctgtacacacgcccgttaacccacttagagggagga 1377
||||| 1327 CGGTGAATCGTTCGCGACCTTGTACACACCGCCGCTACACACCTGAGTGGGAGCA 1386
||||| 1378 ccggaagtggtctgttaccgttaaggaagacgactactaaagctlaaagcgctlaaggg 1437
|||||

```

```

Db 1387 CCCGAAGTGTCTTTGCAACCCGTAAGAGACGAGCTACTAAGGTGAACCTCGTAAGG 1446
Qy 1438 ggtgaagtcgtaacaagttacc 1459
Db 1447 GGTGAAGTGTGAACAGTATGC 1468

RESULT 6
MA5216S 1489 bp DNA BCT 18-FEB-1995
LOCUS Leptospira wellii Worsfold 16S rRNA gene, partial sequence.
DEFINITION U12677.1 GI:558938
ACCESSION U12677.1
VERSION U12677.1
KEYWORDS Leptospira wellii.
SOURCE Leptospira wellii.
ORGANISM Bacteria; Spirochaetales; Leptospiraceae; Leptospira.
REFERENCE 1 (bases 1 to 1489)
AUTHORS Ralph,D. and McClelland,M.
TITLE Phylogenetic evidence for horizontal transfer of an intervening
sequence between species in a spirochete genus
JOURNAL J. Bacteriol. (1994) In press
REFERENCE 2 (bases 1 to 1489)
AUTHORS McClelland,M.
TITLE Direct Submission
JOURNAL Submitted (25-JUL-1994) Michael McClelland, California Institute of
Biological Research, 11099 North Torrey Pines Road, La Jolla, CA
92037, USA

FEATURES
source
1..1489
/organism="Leptospira wellii"
/strain="serogroup Hebdomadis, subserogroup Borincana,
serovar Worsfold1, strain Worsfold"
/isolate="MA52"
/db_xref="taxon:28184"
<1..1489
/product="16S rRNA"
BASE COUNT 383 a 332 c 455 g 316 t 3 others
ORIGIN

Query Match 85.6%; Score 1263.8; DB 1; Length 1489;
Best Local Similarity 94.1%; Pred. No. 3.9e-12;
Matches 1376; Conservative 0; Mismatches 80; Indels 6; Gaps 6;
Qy 1 gatcatgctcagaactaacgcctgagcgcgcttcaaatcgaagtcgagcggttag 60
Db 16 GATCCTGGCTCAGAACTAACCTGCGCGCTTAAACATGCAAGCAAGCGAGTAG 75
Qy 61 caataaccgaagcggaagggtagtaacagt-gtatacttcctcgaatctggagta 119
Db 76 CAATACCTACGCGGCAACGGGTGAGTAACCTGGGTATATCTTCTCGAGCTGGGATA 135
Qy 120 acttccgaagaaagcttaataccggaatgctcgttggatgatacgaattgtagta 179
Db 136 ACTTCCGAAGGGAGAGCTAATAGTATGATGATGATGATGATGATGATGATGATGAT 195
Qy 180 aagattatgcttggaatgagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 239
Db 196 AAGATTATTGCTCGAGCTGAGCCCGCGCGCGGATTAAGTGAAGTGAAGTGAAGTGA 255
Qy 240 accaagcgacagatcgtagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 299
Db 256 ACCAAGCGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 315
Qy 300 cgttcataactcctcgaaggaagcaggttaagaatcttgcctcaatggggaacccctga 359
Db 316 CGGTCCATCT-CTACGGAGGCGACAGTAAAGTCTTGTCAATGGGCGGCAACCTGA 374
Qy 360 agcagcgagcgcgctgaaacgaagaaagcttcgagttgaagttcattaggaac-ggaa 418
Db 375 ACCAGCGAGCGCGGTGAACGATGAAGGCTTTCGATTTGTAAGTTCAATTAAGCAGGGA 434
|||||

```

QY 419 aataagcagcaatgtagatgtagtacctgacctaaagcaccggtactactacgttccagca 478
 |||||
 DB 435 AATAAGCAGCGATGTGATGATGTCTACCTGCTTAAGCACC GGCTTAACCTACGTGCCAGCA 494
 QY 479 gccgcggtatacgtatgtagtcaagcggtgttccgaaatcattggcgctaaagggtgcgt 538
 |||||
 DB 495 GCCGCGGTAACTGATGTGTGCAAGCGTTGTTCCGAATCATTTGGGCTTAAGGGTCCGTA 554
 QY 539 gccgcggtatgtagtcaagcttgaataacgctggcggtaccgggtgacctgtgaac 598
 |||||
 DB 555 GCGGACATGTAAGTCAAGTGTGAATAAGCGGGCTCAACTCCACCTGCACTTGAAC 614
 QY 599 tacaactcggagatttggagagagcaagtggaaatccagctgtagcggtgaatgtag 658
 |||||
 DB 615 TATGTCTCTGGAGTTTGGAGAGCAAGTGAATTCAGGTGACGGGTGAATGGCTAG 674
 QY 659 atacttgaggaagcaccagctggcgaaagcgacttgcgtcctcaaaactgaagctgaagca 718
 |||||
 DB 675 ATATCTGAGAGAAACACCACTGGCGAAGCGACTGCTGGCTTAAACTGACGCTGAGCA 734
 QY 719 cgaagcgtggttagtaagcaaggtatagaatacccgtaataccgctcaaacgttgc 778
 |||||
 DB 735 CGAAGCGTGGGTAGTGAAGCGGATTATACCCCGTATATCCACCCCTAAACGTTGTC 794
 QY 779 taccagttgtggtgttttaaccctcaagtaacgaacctaaagcaatgaagcaagcct 838
 |||||
 DB 795 TACCACTTGTGGGGGTTTAAACCCCTACGTAAGCA -CTAACGATTTAAGTACCGCCT 853
 QY 839 gggagactatgctcgaagagtgaaactcaaggaattgacgggggtcgcgacaagcggtg 898
 |||||
 DB 854 GGGGACTATGCTCGCAAGAGTGAACCTCAAGCAATTGACGGGGGNNMNCACAAAGCGGTG 913
 QY 899 gagcatgtgtttaaattgatgatgatacccaaaactccctgggttgaatgtagctg 958
 |||||
 DB 914 GAGCATGTGCTTAAATTCATGATACGCAAAACCTCACCTAGGCTTGACATGAGAGTGG 973
 QY 959 aatcatgtagaataatagacctcggcgacagatcacaggtgtcgtacgtgtgtcgtca 1018
 |||||
 DB 974 AATTATGTAGATATACATAG -CTTCGGGGCGCTTCACAGTCTCATGTTGTGCTCA 1032
 QY 1019 gctcgtgtcgtgagatgtgtgttgaagtcgcgcgaacgagcgcaacccctatcgtagtctg 1078
 |||||
 DB 1033 GCCTGCTGTGATGTTGGGTTAAGTCCCGCACGACGCGCAACCTTATATGTTG 1092
 QY 1079 cta -ccttaagttgggcactgtgtgacgaactgtcggttgaacaacggagaaagcggtga 1137
 |||||
 DB 1093 CCATCTTTATGTTGGGCACTCGTAAGGAACGCGGCTGACAAACCGGAGAGCGGGGA 1152
 QY 1138 tgaagcacaatccctcaatggccttattgtccagggccacacacgctgtctacaatggtccgata 1197
 |||||
 DB 1153 TGACGTCAAAATCCTCATGGCCTTATGTCTAGGGCAACACAGTGTCTACAATGGCCGTA 1212
 QY 1198 cagaggtgtgcgaactcgcgaagagagactaatctctaaagtctggttcccaagttcgat 1257
 |||||
 DB 1213 CAAAGGGTACCAACCTCGGAGGGGAGGTAACTCAAAAAGCGGCTCCAGTTCCGATT 1272
 QY 1258 ggggtgtcgaactcgcacccaatgaagtcgaatcgtagtaactcgaggttcaagcaagccg 1317
 |||||
 DB 1273 GAGGTGTGCAACTCGACTCATGAATCGGAATCGCTAGTAATCGGGATCAGACATGCCG 1332
 QY 1318 cggtagaatacgttcccggaacttgtaacacacgcccgcgtcaacacacactgaagtgagga 1377
 |||||
 DB 1333 CGGTGATATAGTTCCCGGACCTTGTACACACCGCCGTCACACACCTGAGTGGGAGCA 1392
 QY 1378 ccggaagtggtctttagtaacgttaaggaagcagactactaaggtgaactcgtaaggg 1437
 |||||
 DB 1393 CCGGAAGTGTCTTCCCAACCGTAAGGAAGACGACTACTAAGGTAAACCTGTGAAGGG 1452
 QY 1438 ggtgaagctgtaacaaggtacc 1459
 |||||
 DB 1453 GGTGAAGTGTGTAACAAGTATGAC 1474

RESULT 7
 LB012670 1513 bp DNA BCT 18-FEB-1995
 LOCUS
 DEFINITION Leptospira borgpetersenii Hardjo bovis/sponselee 16S rRNA gene,
 partial sequence.
 ACCESSION U12670
 VERSION U12670.1 GI:558929
 KEYWORDS
 SOURCE Leptospira borgpetersenii.
 ORGANISM Bacteria; Spirochaetales; Leptospiraceae; Leptospira.
 REFERENCE 1 (bases 1 to 1513)
 AUTHORS Ralph,D. and McClelland,M.
 TITLE Phylogenetic evidence for horizontal transfer of an intervening
 sequence between species in a spirochete genus
 JOURNAL J. Bacteriol. (1994) In press
 REFERENCE 2 (bases 1 to 1513)
 AUTHORS McClelland,M.
 TITLE Direct Submission
 JOURNAL Submitted (25-JUL-1994) Michael McClelland, California Institute of
 Biological Research, 11099 North Torrey Pines Road, La Jolla, CA
 92037, USA
 FEATURES
 source Location/Qualifiers
 1..1513
 /organism="Leptospira borgpetersenii"
 /strain="serovar hardjo, strain Hardjo bovis/Sponselee"
 /isolate="BD50"
 /db_xref="taxon:174"
 <1..1513
 /product="16S rRNA"
 BASE COUNT 385 a 346 c 464 g 315 t 3 others
 ORIGIN
 Query Match 85.5%; Score 1262.2; DB 1; Length 1513;
 Best Local Similarity 94.0%; Pred. No. 4,1e-12;
 Matches 1375; Conservative 0; Mismatches 81; Indels 6; Gaps 6;
 QY 1 gatcagtgctgaagactaagcgtggtggcggtcttaacaatgcaagtgtaggggtag 60
 DB 40 GATCTGTGCTCAGAACTAACGCTGGCGGCGCTTTAAACATGCAAGCGGAGTGG 99
 QY 61 caatacttagcggcgaagcggtgtagtaacagt-ggtaactctccctcgaagcttgagata 119
 DB 100 CATACTCAGCGGCGCAACGGGTGATGTAACACGTGGTAACTCTCCCGAGTGGGATA 159
 QY 120 actcttcgaagaagaaactaatacggatagctcgtgtgtagtaacaagattgatagta 179
 DB 160 ACTTTCGGAAGGGGAGCTAACTGTGATAGTCCCGAGAGGTCAATGATTTTTCGGGTA 219
 QY 180 aagattattgtctgtagagtagcccgcggtacgtagctagttgtaggtgaatgtctc 239
 DB 220 AAGATTATTGTCTCGCGCTGAGCCCGCCCGCATTAAGCTAGTGGTGAAGTAATGGCTC 279
 QY 240 accaagcgacagatcgttagcggcgttgagaggtgtccggccacaaatgaaactgagaca 299
 DB 280 ACCAAGCGACGATCGTGTAGCGCGCTGAGAGGCTGTGGCCACAATGAAGTAAGTGAACA 339
 QY 300 cggctcatatctcctaagggagcagcttaagaatctgtctcaataggggaaacccctga 359
 DB 340 CGGTTCATACT-CTACGGGAGGCGACAGTTAAAGAACTTGTCTCAATGGGGGAACCTTA 398
 QY 360 agcagcgacccgcgtgaaagaaaggtcctcgagttgtaagttcatatgaagca-gga 418
 DB 399 AGCAGCGACCCCGCTGAACGATGAAGGCTTGTGGAATTGTAAAGTTCAATAGCAGAGAA 458
 QY 419 aataagcagcaatgtagatgtagtacctgacctaaagcaccggtactactacgttccagca 478
 DB 459 AATAAGCAGCGATGTGATGATGTGTACCTGAAGCACC GGCTTAACCTACGTGCCAGCA 518
 QY 479 gccgcggtatacgtatgtagtcaagcggtgttccgaaatcattggcgctaaagggtgcgt 538
 DB 519 GCCGCGGTAACTGATGTGTGCAAGCGTTGTTCCGAATCATTTGGGCTTAAGGGTCCGTA 578

```

QY 539 ggcggattctgaatcagctgctgaaacatcgccggtcgaacccgtgacctgacctgaac 598
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 579 GGGGACATGTAATCTCAGGTGTAAGAACTCGGGCTCACTCCGAGCGCTCAGCTTGAAC 638
QY 599 tacaacgtcgagtttgagagagcgagctggaattccagctgtgagcggtgaatcgtag 658
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 639 TATGTGCTGAGTGTGGAGAGGCAAGTGAATTCAGGTGTAGCGCGTAATATGCGTAG 698
QY 659 atactctggagaaacacccagctggcggaagcgactgtcgtctcaaacctgacgtcga 718
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 699 ATATCTGGAGAAACCAACAGTGGCCAAAGCCACTTGTGCTTAATAACTGAAGCGTAGACA 758
QY 719 cgaagcgtgggttagtaaacgagatagatacccggtlaatcaacgacctaaacgtgtgc 778
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 759 CGAAAGCGTGGGTAGTGAACGGATTTAGATACCCGGTAAATCCAGCCCTAAAGCTTGTCT 818
QY 779 taacagttgttgggtgtttaaaccctcagtaacgaacctaagctgaagtagacgcct 838
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 819 TACCAAGTGTGGGGGTTTAAACCTCAGTAACCAH-CTAAACGATTAAGTAGACCCGCT 877
QY 839 ggggactatgctcgcaagagctgaaactcaagaagattgacggtgggtccgcaagcggtg 898
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 878 GGGGACTATGCTCCGACAGAGTAAGAACTCAAGGAATTGACGGGGTNNMCAAGCGGTG 937
QY 899 gagcaatgtttaaattcagatgatacccaaaaacactcactggtggttgaatgactgtg 958
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 938 GAGCATGTGGTTTAATTTGATGATATGCGCAAAACCTCACTAGGCTTGACATGGAGTGG 997
QY 959 aatcatgtagaagataatagagccttcggcgcaagatcacaggtgtcgtacatgtgtcgtca 1018
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 998 AATCATGTAGATATCATGAG-CTTCGGGGCGGCTTACAGGAGTGTGATGGTGTGCTGCA 1056
QY 1019 gctcgtgtcgtgagatgtgtggttgaagctccgcaagcgcgcaacccctacgtatgtgtg 1078
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1057 GCTCGTGTGTAGATGTGGGTAAAGTCCCGCAACGAGCCGACACCTCATTTATGTTG 1116
QY 1079 cta-ccttaagtttgagcactgtgtacgaacatcggttgacaacccggaagcggtgga 1137
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1117 CCATCTTTCAGTTGGGCACTCTTAAGAACTCCGGTGTGCAAAACCGGAGAAAGCGGGGA 1176
QY 1138 tgaagctcaaatccctacatgagccttataatgctcgaagcgcaacacgtgtctacatgtgcgata 1197
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1177 TGACGTAATATCTCTATGSCCTTTATGTGTAGGGCAACACAGTGTCTACAAAGCGCGGTA 1236
QY 1198 cagaggtgtcgcaacatcgcaagagagagtaatctcctaagctgtgtccagttcgatc 1257
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1237 CAAGAAGTGAAGCCAACTCGAGAGGGGAGCTAATCTCAAAAAGCCGGTCCCACTTGGATT 1296
QY 1258 ggggtctgcaactcgaccccaatgaatcggaatcgctagtaatcgcgatcagcatgctcg 1317
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1297 GGAAGTCTGCAACTCGAAGTCACTGAAGTCCGAAATCGCTAGTAATCGCGGATCAGCATGCCG 1356
QY 1318 cgggtgaatacgttcccggaactgtgtacacacgcccgttcaacacactgaagtgggagaga 1377
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1357 CGGTGAATAGCTTCCCGGAGCTTGTACACACCCCGGTACACACACTTAAGTGGAGAGA 1416
QY 1378 ccggaaggtgtcttgttgaacgtaagagagacagactactaaggtgaactcgtlaaggg 1437
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1417 CCGGAAGTGTCTTTGGCAACCCCAAGGAAGACAGACTACTAAGGTAAACTCTGTGAAGGG 1476
QY 1438 ggtgaagcttaacaaggtlac 1459
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1477 GGTGAAGTCTTAACAAGTAGC 1498

```

```

RESULT 8
LOCUS LB012669 1489 bp DNA BCT 18-FEB-1995
DEFINITION Leptospira borgpetersenii 1627 Burgos 16S rRNA gene, partial
sequence.
ACCESSION U12669
VERSION U12669.1 GI:558928

```

```

KEYWORDS
SOURCE Leptospira borgpetersenii.
ORGANISM Leptospira borgpetersenii.
Bacteria; Spirochaetales; Leptospiraceae; Leptospira.
REFERENCE
AUTHORS 1 (bases 1 to 1489)
TITLE Ralph, D. and McClelland, M.
JOURNAL Phylogenetic evidence for horizontal transfer of an intervening
          sequence between species in a spirochete genus
          J. Bacteriol. 176 (19), 5982-5987 (1994)
MEDLINE 95014031
REFERENCE 2 (bases 1 to 1489)
AUTHORS McClelland, M.
TITLE Direct Submission
JOURNAL Submitted (25-JUL-1994) Michael McClelland, California Institute of
          Biological Research, 11099 North Torrey Pines Road, La Jolla, CA
          92037, USA
FEATURES
    source
        1..1489
            /organism="Leptospira borgpetersenii"
            /strain="serovar balcanica, strain 1627 Burgos"
            /isolate="BD30"
            /db_xref="taxon:174"
            <1..1489
            /product="16S rRNA"
BASE COUNT 385 a 330 c 454 g 311 t 9 others
ORIGIN
Query Match 85.4%; Score 1261; DB 1; Length 1489;
Best Local Similarity 93.8%; Pred. No. 4,3e-12;
Matches 1372; Conservative 0; Mismatches 84; Indels 6; Gaps 6;
QY 1 gatcatgctcgaactaacgcttgcgcgcgtctttaaacaatcgaagctcagcggtgag 60
Db 16 GATCTGCTCAGAACATACAGCTGCGCGCTTAAACATCAACATCAAGCGGAGTAG 75
QY 61 caataactagcgagcaacggtgtgagtaacagt-ggtaactcttcctcgcagcttgagata 119
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 76 CAATATCTAGCGGCAACGGGTGATACAGTGAGTGAATCTTCTCCGAGTCTGGGATA 135
QY 120 acttccgaagaaagcctaataccgagatagctcgttcttgatcaacaagattgtatagta 179
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 136 ACTTTCGAAAGGGAGACTAATACCTGATGATCCGAGAGCTATACGATTTTTCGGGTA 195
QY 180 aagattatctcttgagagatgagcccgcgcgatagctagttggtgaggtaaatgctc 239
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 196 AAGATTATTGCTCGAGATGAGCCCGCGCGATTAGCTAGTGTGAGGTAAATGCTC 255
QY 240 accaagcgagacatcgtgtagcgcgccttagagaggtgtccgcgcacaaatggaactgagaca 299
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 256 ACCAAGCGGAGATCGTAGCCCGGCTGAGAGGGTTCGGCCACAAATGMACTGAGACA 315
QY 300 cgttccatactctcactcagggagcgagcttaagaatctctcctaattgggggaaaccttga 359
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 316 CGGTCCATACCT-CTACGGGAGGACAGCACTTAAGATCTTGTCAATGGGGGAACCTTGA 374
QY 360 agcagcgagcgcggtgaaagaaaggtcttcgattgttaadtctcatagga-99aa 418
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 375 AGCAGCGAGCGCGCTGAGACGATGAAGGCTTTCGGATTGTAAAGTTCAATAGCAGGAA 434
QY 419 aaataagcagcaatgtgtatgtatgtactgtcctaagaacacggcttaactacgttgcaca 478
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 435 AATATAGCAAGCATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 494
QY 479 gccgcgtlaatacgtatgtgcaagcgtgtgtcggaaatcattgggcgttaaggggtgctga 538
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 495 GCGCGGTAAATACGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 554
QY 539 ggcggattgtgtaagcaggtgtgaaacatcggtggttcgaacccgtgtgacctcgtgaac 598
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 555 GCGGACATGTAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 614
QY 599 tacaagctcgtgagtttgagagagcgcaagtggaattccaggtgtgacggtgtgaaatgcgtag 658

```

```

Db      615 TATATCTGTGAGATTGGAGAGGCAAGTGAATTCAGGTGAGGGGTGAATGCGTAG
Qy      659 atatctggaggaacacacgaatggcgaagcgactgtctgctcaaaactgaagctgaagca
Db      675 ATATCTGGAGGAACACCACTGGCGAAGGCGACTGCTGGCTTAATAAATCAGCTGAGGCA
Qy      719 cgaagcgggtagtaagaacgagatagaatacccggtatccacgacctaaactgtctc
Db      735 CCAAGCGGTGAGTGAACGGGATTAGTATACCCCGTATCCAGCCCTTAACGTTGTC
Qy      779 taccagttgtgggggttttaaccctcagtaacgaacctaagagtaagtagaccgct
Db      795 TACCACTTTGTTGGGGGTTTAAACCTCACTAAGCAAA-CTAACGATTAACTAGACCGCT
Qy      839 ggggactatgctcgaaggtgaacactcaagaagaatgtgacgggggtccgcacaaagcggt
Db      854 GGGGACTATGCTCGCAAGAGTGAACCTCAAGGAATTGACGGGGTNNNCACAAAGCGGTG
Qy      899 gagcatgtgttaatttcgatgataccccaaaacctcaacctggggtcttgacatgactcg
Db      914 GAGCATGTGTTTAATTCATGATACGCCAAAACCTCACCTAGGCTTGACATGGAGTGG
Qy      959 aatcatgtagatatataatagaccttcgggcagatcacaagttgcatggtgtgtcga
Db      974 AATCATGTAGATACATAGAG-CTTCGGGCGCTTCACAGTGTGCTCATGTTGTCGTCA
Qy      1019 gctcgtgtgtgagatagttgggttaagttcccgcaagacgcaacccctcatgtaagttg
Db      1033 GCTCGTGTGTGATGATGTTGGGTTAAGTCCCGCAACGACGCGCAACCTCATGTTAAGTTG
Qy      1079 cta-ccttaagttggaactgtgtgacgaacatgcggtgacaaacccggaagagcgagga
Db      1093 CCATCATTTAGTTGGGACTCTCGTAAGGAATCGCGGTGACAAATCGAGAGAGCGGGCA
Qy      1138 tgaagtcgaacccctcaatggccttattgtccaggggcacacacgltgataagtcgagta
Db      1153 TCACGTCANNNNNNATGCGCTTTATGTAGGGCAACACAGCTGCTACATGCGCGGTA
Qy      1198 cgaaggtgtcgaactcgaagaagaggaagatctcttaaaagtcggttccagttgactg
Db      1213 CAAGAGGTAGCCCACTCGCGAGGGGAGCTAATCTCAAAAAGCGGTCCAGTTCGATTT
Qy      1258 ggggtgtcgaactcgaaccccatgaagaatcggaatcgctagtaatcgggagtcagcagcg
Db      1273 GGAAGTGTGCACTCGACTCGATGAAGTCCGAATCGCTAGTAATCGGGGATCAGCAATGCCG
Qy      1318 cgggtgaatcgttcccggaacctgtgacacacgcccgtcacaccactgagttggagaga
Db      1333 CGGTGAATAGCTTCCCGGACCTTGTACACACGCCCGCTCACACACCTGAGTGGGAGCA
Qy      1378 ccggaagtgtcttgttaccgtaangagagacgaactactaagttgaanaactcgttaagg
Db      1393 CCGAAGTGTGTTTGGCCACCGCAAGGAACGACTACTAAGGTAAACTCTCTGAAAGG
Qy      1438 ggtgaagtcgtaacaagtlacc 1459
Db      1453 GGTGAAGTGTGAACAAGTATGC 1474

```

```

RESULT 9
LOCUS   LN012671      1516 bp      DNA          BCT      18-FEB-1995
DEFINITION   Leptospira noguchii Fort Bragg 16S rRNA gene, partial sequence.
ACCESSION   U12671
VERSION     U12671.1 GI:558930
KEYWORDS
SOURCE
ORGANISM   .
            Leptospira noguchii.
            Bacteria; Spirochaetales; Leptospiraceae; Leptospira.
REFERENCE   1 (bases 1 to 1516)
            Ralph, D. and McClelland, M.
            Phylogenetic evidence for horizontal transfer of an intervening
            TITLE

```

```

sequence between species in a spirochete genus
JOURNAL    J. Bacteriol. 176 (19), 5982-5987 (1994)
MEDLINE    95014031
REFERENCE   2 (bases 1 to 1516)
AUTHORS    McClelland, M.
TITLE      Direct Submission
JOURNAL    Submitted (25-JUL-1994) Michael McClelland, California Institute of
            Biological Research, 11099 North Torrey Pines Road, La Jolla, CA
            92037, USA
FEATURES
    source          Location/Qualifiers
    1..1516
    /organism="Leptospira noguchii"
    /strain="serogroup Autumnalis, subserogroup Fort Bragg,
    serovar fortbragg, strain Fort Bragg"
    /isolate="NB36"
    /db_xref="taxon:28182"
    <1..1516
    /product="16S rRNA"
    BASE COUNT      392 a      342 c      462 g      314 t      6 others
    ORIGIN
    1 gatcatggtctgaagaactaagcgtggcggtcttaaacatgcaagtcgagcggtgag 60
    Db      43 gatcttgctcgaactaacgctggcgcgctttaaacaatgcaagtcgaagcgatg 102
    Qy      61 caataccttagcggcgaacgggtgagtaaacgt-aggtaactctcccgagtcgtgagata 119
    Db      103 CATACTCAGCGCGGCAACGGGTGATGATACACGTTGATATCTTCCTTCGAGTGGGATA 162
    Qy      120 acttccgaagaagaactaatacggatagtcgtgtgatacagaagattgataagta 179
    Db      163 ACTTTCGAAGGGAAGCTAATCTGTGATATGCCGAGATCATTAAGATTTTTCGCGTA 222
    Qy      180 aagattattgtcttgagatgaagccggcgcgagtagtaagttgtgaggttaagtcgc 239
    Db      223 AAGATTCAATTGCTCGGAGATGAGCCCGCTCCGATTAGCTAGTTGGTAAAGTAAAGGCTT 282
    Qy      240 accaagcgacgaatcgtgtgacggccttgagagggtgtccggcacaaatgaaactgagaca 299
    Db      283 ACCAAGCGCGCATGCTGATGCGCGCTGAGAGGGTGTTCGCCCAATGAACTGAGACA 342
    Qy      300 cgggtcaatactcctacgggaagcagatgaagaatctgtcctcaatgggggaacctgga 359
    Db      343 CGGTCCATACT-CTACGGGAGGCGCAGGATTAAAGAACTTGTCAATGGGGGAAACCTGA 401
    Qy      360 agcagcgacggcggtgaacgaagaggtcttcgggttgaagttcatltagca-ggaa 418
    Db      402 AGCAGCGACGCGCGTGAACGATGAAGGCTTCGATTGAAGTTCAATFACACAGGGA 461
    Qy      419 aaataagcagcaatgtgatactgtaacctgctaaagcaccggtctaactacgttgcagca 478
    Db      462 AATATAGCAACATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 521
    Qy      479 gccggtgaatacgtatggtgcaagcgtgtgtcggaatcaatgtggcgtaaaagggtgctga 538
    Db      522 GCCGCGTAAATACGTATGTCGCAAGCGTTGTCGGAATCATTTGGGCGTAAAGGTCCTGA 581
    Qy      539 ggcggatctgttaagtagtgaataacggcggtcccaaccggtggccggtgacttgaac 598
    Db      582 GCGGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 641
    Qy      599 tacaagtcgtgaagtttgggaagagcaagtggaattccaggtgtgagcggtgaatgtgtag 658
    Db      642 TATGTCTCTGGAATTTGGGAGAGCAAGTGAATTCAGAGTGTGAGCGGTGAATATCGTAG 701
    Qy      659 atatctggaggaacacacgaatggtgaaagcgactgtgtgtcctcaaaactgaagctgagaga 718
    Db      702 ATATCTGGAGGAACACCACTGGCGAAGGCGACTTGTGCTGCTTAATAACTGACGCTGAGGCA 761

```


QY 719 cgaagcggtggaagtaaacgaggtatagatacccggtatccacgcccctaacggtgtc 778
DB 762 CGAAAGCGGTGGATAGTAACGGGATAGATNCCCGGTAATCCAGCGCCCTTAACGGTTGTG 821
QY 779 tacaagtttttgggggttttaaccctcaatgaacgaacctaacggtatgaacgcct 838
DB 822 TACCAAGTTGGGGGTTTAAACCTTCACTAGACGA-CTAACGGATTAAGTAGACCGCCT 880
QY 839 ggggaactatcgcgcaagaatgaactcaagaagaattgaacgggggtccgcaagcggtg 898
DB 881 GGGGACATATGCTCCAGAGAGTAACCAAGCAATGACGGGGGTNNNCAACAAGCGGTG 940
QY 899 gaggatgtgttaattcgtatgaatacccaaaaaccccaactggtgttgatgaatgactg 958
DB 941 GAGCATGGGTTTAAATGAGATACCGAATAACCTACCTAGGCTTGAATGGAAGTGG 1000
QY 959 aatacgtatagatatatgaaccttcgggacagatccaaggtgctgcatgtgtcgtca 1018
DB 1001 AATCATGTAGAGATACATGAG-CTTCGGGGCGCTTACAGGTGCTGCATGGTTGCTGCA 1059
QY 1019 gctcgtcgtgaagatgttgggttaagtcccgcaacgagcgcaacccctatcgtatgtg 1078
DB 1060 GCTGGTGTGATGATGTTGGGTTAAGTCCGCAACGAGCGCAACCTCACCTTATGTNG 1119
QY 1079 cta-ccctaaagtggacatgtgtacgaactgcccgtgaacaaccggaggaagcgagga 1137
DB 1120 CCATCATTCAGTTGGGCACTGTAAGAACTGCGGTGACAAACCGAGAAAGCGGGGA 1179
QY 1138 tgaactcaaatccctcatgtcttatagtccagggccacacacagtcgtacaaatggccgata 1197
DB 1180 TGAGCTCAATTCCTCATGAGCCCTTATGTCFAGGGCAACACAGTCFCAATGGCGGGTA 1239
QY 1198 cagaaggtcgcgaactcgcaagggaggaactatccttaaaagtgcgtccagttcgagt 1257
DB 1240 CAAGGGTAGCCAACTCCGAGGGGAGCTATATCTCAAAAAGCCGGTCCCGTGGGAT 1299
QY 1258 ggggtcgaactcgaccctcaatgaatgcgaatcgtaactcgagatcaagcatgctcg 1317
DB 1300 GAGCTCTGCAACTGCATCTCAGTAGAGTGGAAATCGCTAGTAATCGCGATCAGCATGCGG 1359
QY 1318 cgggtgaatacgttcccggaacctgtacacacccgcttcaacacacactgagtggagaa 1377
DB 1360 CGGTGAATACGTTCCCGGACCTTGTACACACCGCCGTACACACCTGAGTGGGAGACA 1419
QY 1378 cccgaagtggtcttgttaacgtaagagagacagactactaagtggaactcgttaagg 1437
DB 1420 CCGGAGGTGGTCTTGGCCAAACCGCAAGAGCAAGCACTACTAAGTGAACCTCGTAAGGG 1479
QY 1438 ggtgaagtcgtacaaggtacc 1459
DB 1480 GGTGAAGTGTAAACAAGTACG 1501

RESULT 10
LB16SRDNP 1425 bp DNA 02-DEC-1993
LOCUS LB16SRDNP 1425 bp DNA 02-DEC-1993
DEFINITION L. borgpetersenii gene for 16S ribosomal RNA (partial).
ACCESSION Z21630.1 GI:433579
VERSION 221630.1 GI:433579
KEYWORDS 16S ribosomal RNA.
SOURCE Leptospira borgpetersenii.
ORGANISM Leptospira borgpetersenii.
Bacteria; Spirochaetales; Leptospiraceae; Leptospira.
REFERENCE
1 (bases 1 to 1425)
Hookey, J.V.
Phylogeny of Leptospiraceae and related Spirochaetes
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 1425)
Hookey, J.V.
REFERENCE Direct Submission
AUTHORS Submitted (09-FEB-1993) Hookey J.V., Public Health Laboratory
JOURNAL Service, Leptospira Reference Laboratory, Stonebow Road, Hereford,

Herefordshire, United Kingdom, HRI 2BR
FEATURES
Location/Qualifiers
source 1.1425
/organism="Leptospira borgpetersenii"
/strain="JAVANICA VELDRAT BATAVIA 46"
/db_xref="taxon:1174"
rRNA
1.1425
/product="16S ribosomal RNA"
BASE COUNT 369 a 327 c 430 g 297 t 2 others
ORIGIN

Query Match 84.9%; Score 1254.6; DB 1; Length 1425;
Best Local Similarity 94.2%; Pred. No. 5,6e-12;
Matches 1343; Conservative 2; Mismatches 76; Indels 4; Gaps 4;

QY 16 ctaacgctggcgcggtcttcaacatgaagtc-gagcggggttaacatacctacgagc 74
DB 1 STARCGCTGGCGGGCGCTTTAAACATCCAAATCCAAAGCGGAGTACCAATCTACAGCGG 60
QY 75 gaaaggtgagtaacagct-gttaatcttcctccgaagctcgtgataacttccgaagga 133
DB 61 GAACGGGTGAGTAAACAGTGGATATCTTCTCGAGTCTGGGATTAATCTCCGAAGAGG 120
QY 134 aagctaaacggaatagctcgtgtgtatcaacaagaattgataagtaagaattatgct 193
DB 121 GAGCTAATACGTGATAGTCCCGAGAGGTATAGATTTTTCGGTTAAAGATTATGCTC 180
QY 194 ggaataagcccgagcgccgattagcttagtgtgaggttaatgtgtcaccgaagcgacat 253
DB 181 GGAATAGACCCCGCGCCGATTAAGTGTGTAAGTAATGCTTCACAAAGCGAGCGAGAT 240
QY 254 cggtagccgacctgagaggggtgtccggccacaatgtaactgagaacagcgatcactcct 313
DB 241 CGGTAGCCGGGCTGAGAGGGGTTCGGGCCCAATGAACTGAGACAGGTCATACCT 300
QY 314 acggagggcagcaatgaagaatcttgcataagggggaacccctgaagcagcgagcgcg 373
DB 301 ACGGAGGACCACTTAAGTAATCTTGTCAATGGGGGGAACCTGAAAGCAAGCAAGCGCCG 360
QY 374 gtgaacgaagaaggtcttcgattgtataaagttcataagca-ggaaaaaataagcagcaat 432
DB 361 GTGAACGATGAAGGCTTTCGGATTTGTAAGTCAATTAAGCAGGGAATAATGACAGCGAT 420
QY 433 gtgatgtatgtaactcgtcctaaagcagcggtcaactaogtgcagcagcgcggttaacg 492
DB 421 GTGATGATGTACTGCTTAAGCAACCGGCTAAGTACGTGCACAGCGCGGTAAATACG 480
QY 493 tatgttgaaagcgttgcgtcgaatacatgtggcgtlaaaggtgctgtagcgagattgtlaag 552
DB 481 TATGGTCAAGCGTTGTTCCGAATCATTTGGCGTAAAGGTGCGTAGCGGACATGTAAAG 540
QY 553 tcaagtgtaaaactcgtcggtcacaaccgtgtgcctgcactgaaactgaacatcgtgaagt 612
DB 541 TCAGGTGTGAATACTGCGGGCTCAACTCGACGCTGCACTTGAAACATATGCTTGGAGT 600
QY 613 ttgggaagggcgaagtgtgaattccaggtgttagcgtgtgaattcgtatgataatctggagaa 672
DB 601 TTGGGAGAGGCAAGTGAATTCAGGTGTAGCGGGAATCGTAGATATATCTGAGAGAAC 660
QY 673 accagtgaggaagcgacttcgtgctcaaaactgaacgctgagcgacgcaagcggtggtla 732
DB 661 ACCAATGCGGAAGGGGACTGTGCGCTAAAACTAGCCTGAGGACGCAAGCGGTGTA 720
QY 733 gtaaacggatagatatacccggttaataccagccttaaaagttgtctacaggtgttgg 792
DB 721 GTGAACGGGATTAAGTATACCCCGTAAATCCAGCCCTTAACGTTGCTCAAGTGTGGG 780
QY 793 ggttaacccctcaatacgaaccccaacggaattagtagacccgctggggactatgctcg 852
DB 781 GGTTTTAACCTCAGTAAAGAACTTAACGATTAAGTAGACCGCTGGGACTATGCTCG 840
QY 853 caaggtgaactcaagaagaattgacggtggtcgcacaaagcggtgagcatgtgttla 912


```

|||||
Db 841 CAGAGTGAACCTCAAGGAATTGACGGGGTCCGCACAAAGCGGTGAGACATGTGGTTA 900
Qy 913 attcatgtatccccaaaactcaactccttggttgatgacatgatactgaatcgttagat 972
Db 901 ATTGCGTGTAGTACGGCGAACTCACCCTAGGCTGTGACATGAGTGAATCATGTAGTAAAGAT 960
Qy 973 atatgagcttcgggacagatcaccaagtgctcatgtgtctgcagactcgtgtctgag 1032
Db 961 ACATGAGCCTTGGGGCGCTTACAGAGTGTGCTCATGTGTTGCTGCTACGTCGTGTGAG 1020
Qy 1033 atgttggttaagctcccgcaagcgcaaccctcactcgtatgttcta-ccctaaagttg 1091
Db 1021 ATGTTGGTAAATGTCGCCGACAGAGCGCAACCTCATTATGTTGCCATTCATTCAGTTG 1080
Qy 1092 ggcactggtgacgaactcgcgcgttgacaacacggaggaagcggggagatgcacgtcaact 1151
Db 1081 GGCACCTCGTAAGAACGTGCGGTGACAAACCGAGAGAGCGGGGATGACGTCAATTCCT 1140
Qy 1152 catgagccttatgtccagggccacacacgctgtctacaatgtgcacatacagaaggtgcacaa 1211
Db 1141 CATGGCTTTATGTCTAGAGGCAACACAGCTGCTACATAGGCCGCTACAAAGGTTAGCCAA 1200
Qy 1212 ctcgcaagcgagagactaactctctaaagtcggttccagttcgatgtgggtctgcaactc 1271
Db 1201 CTCGCGAGGGGAGGTAACTCAAAAAGCGGTCAGTTCCGATTGGAGTGTGCAACTC 1260
Qy 1272 gaccctatgaatcggaatcgcgttaagtaactcgagatcaacatgcccgggtgataagttc 1331
Db 1261 GACTCCATTAAGTCGGAATCGCTAGTAATCGCGGATCAGCATGCCCCGGTGAATACGTTTC 1320
Qy 1332 ccggaactctgtacacacgcgcctcaccacacctgagtgaggagacccgaagtgtctt 1391
Db 1321 CCGGACTGTTATGACACGCCCTCTACACCACCTGAGTGGGAGCAACCCGAATGGTCTT 1380
Qy 1392 tcttaacgtaaggagacagactactaagtgaaactcgtlaaag 1436
Db 1381 TCCCAACCCGCAAGGAGACAGACTACTAAGGTGAATCGTGAAG 1425

RESULT 11
WB4616S 1486 bp DNA BCT 18-FEB-1995
LOCUS Leptospira weilii Sarmin 16S rRNA gene, partial sequence.
DEFINITION U12673
ACCESSION U12673
VERSION 012673.1 GI:558932
KEYWORDS
SOURCE
ORGANISM
Leptospira weilii.
Bacteria; Spirochaetales; Leptospiraceae; Leptospira.
REFERENCE
1 (bases 1 to 1480)
AUTHORS Ralph D. and McClelland, M.
TITLE Phylogenetic evidence for horizontal transfer of an intervening
sequence between species in a spirochete genus
JOURNAL J. Bacteriol. 176 (19), 5982-5987 (1994)
MEDLINE 95014031
REFERENCE 2 (bases 1 to 1486)
AUTHORS McClelland, M.
TITLE Direct Submission
JOURNAL Submitted (25-JUL-1994) Michael McClelland, California Institute of
Biological Research, 11099 North Torrey Pines Road, La Jolla, CA
92037, USA

FEATURES
source
1..1486
/organism="Leptospira weilii"
/strain="serogroup Icterohaemorrhagiae, subserogroup
Sarmin, serovar sarmin, strain Sarmin"
/isolate="WB46"
/db_xref="taxon:28184"
<1..1486
/product="16S rRNA"
BASE COUNT 384 a 332 c 452 g 315 t 3 others
ORIGIN

```

```

Query Match      84.8%: Score 1251.8; DB 1; Length 1486;
Best Local Similarity 94.0%: Pred. NO. 5.9e-12;
Matches 13/5: Conservative 0; Mismatches 80; Indels 7; Gaps 7;

Qy 1 gacatggtcgaacaaacacgctgagcggtcgtcttaacaatgcaagtcgagcggtag 60
Db 14 GATCCCTGCTCAAGACTAAGCTGGGGCGCGCTTAAACATGCAAGTCAACGGGATG 73
Qy 61 caatactagcggcgaaagcggtgagtaacagt-aggtaactcctccgagtcgtgata 119
Db 74 CAATATCCAGCGCGCAACCGGTGAGTAACACGCTGGGTATCTCCCGCAGCTCGGATA 133
Qy 120 acttccgaaagaaagcctaataccgataagtcctgttgatcacaaagttgataagta 179
Db 134 ACTTCCGAAGGGAGCTTAATCTGATATGCCGAGAGTCAATATGATTTTGGGTA 193
Qy 180 aagattatctgttgagatgagcccgcgccgattagtagttgtgaggtatggtc 239
Db 194 AAGATTATTGCTCGGAGCTGAGCCCGCCGCGATTAGCTAGTGGTGAAGTAATGGCTC 253
Qy 240 accaaggcgaacgactgtagcggcctgagaggtgtccgcccacaa tgaactgaaca 299
Db 254 ACCAAGCGCAGCATCGTAGCGGCTGTAGAGAGGTGTTGCGCCACAAATGAACCTGAGACA 313
Qy 300 cgttcatactccttaaggaggaagcagatgtaagaactctgtcgaatggggggaacctga 359
Db 314 CGGTCCATACT-CTACGGAGGAGCAGCATTAAGATCTTGCTCAATGGGGGAAACCTGA 372
Qy 360 agcagcgacgcgcgctgaaacgaagaagctcctgatactgataagttcatlaagca-9gaa 418
Db 373 AGCAGGACGCGCGCTGAAACGATGAAGGCTTCGGATTGAAAGTTCAATGACAGAGAA 432
Qy 419 aaataagcagcaatgtgatgatagttaactcgtcctaaagcaccggtcactaactgtgcaga 478
Db 433 AATAAGCAGCATGTGATGATGTGTACTGCTCAATTAATCAACCGCTTAACGTGCAGACA 492
Qy 479 ggcgcgtaactcgtatgtgtgcaagcgtgtgttcggaataacttgggggtaaaggttcgta 538
Db 493 GCCGCGTAAATGCTATGTGTGCAAGCGTGTTCGGAAATATTGGGCTTAAGGGTCCGTA 552
Qy 539 ggcggaatttgaagtcaggtgtgtaaaactcgcggctcaaacccgtgagctgcacttgaac 598
Db 553 GCGGACATGTAAGTCAGGTGTC-AAAACGCGGCTCAACCTGCACACTGCACCTTGAAC 611
Qy 599 tacaagtcctggagcttctggagaaggaagtggaaattccagatgtgtagcgtgaaatgtcgtag 658
Db 612 TATGTGCTGTGAGATTGGAGAGGCAAGTGAATTCACAGGTGTAGCGGTGAATGCGTGA 671
Qy 659 atactgaggaagcaaccagctggcggaagcgactgtgtgctcaaaactagactgaagca 718
Db 672 ATTCTGGAAGAACACCACTAGTGGGAGGCGACTGTGGCTTAACCTGACGTGAGGCA 731
Qy 719 cgaagcgttggtagtaaaacggaattagatlaaccccgtaatccacgcctaaacgtttgc 778
Db 732 CGAAACGCTGGGTAGTGAACGGGATAGATACCCCGTATTCACCCCTAAACGTTGTC 791
Qy 779 taccagttgtggggttttaaccctcagtaagcaactgaagcattgaagttagacgcgct 838
Db 792 TACCACTGTGGGGGTTTTAACCTCAGTAACGAA-CTAACGGATTAAAGTAACCGCCT 850
Qy 839 gggactatgctcgcgaagtgaaactcaaaagaaattgaacgggggtcgcgcacaagcggtc 898
Db 851 GGGGACTATGCTCGCAAGAGTGAAGTCAAAAGCAATTGACGGGGGNNNNACACAAGGGTG 910
Qy 899 gacactgtgttaactcgtatgataccccaaaaaactcaactcgtgggtctgacatgtatcgt 958
Db 911 GACCATGTGGTTAAATTGATGATGATACGCAAAAACCTCACTAGCTTGAATGATGAGTGG 970
Qy 959 aatcatgtgagatatagagccttcggcgagatcaacaggtgtctgacatggtgtgtcgtca 1018
Db 971 AATTATGTAGATATACATGAG-CTTCGGGCGCGCTTACAGAGTGTCTCATGTTGCTCA 1029

```


|||||
Db 1141 TCATGACCTTTATGCTAGTACGGAACACACGCTGTACATGCGCGGTACAAAGGATGCCA 1200
Oy 1211 actcgcaagaaggagtaactctcctaagaagtcggtccagatcggtatgggtcgtgaact 1270
Db 1201 ACTCGGAGGAGGAGGAGTATCTCAAAATCCGTCACGTCGATTCGATTCGATTCGATTC 1260
Oy 1271 cgaaccacgaagtcggaatcgctatgaatcgatcgatcgatcgatcgatcgatcgatcgat 1330
Db 1261 CGACTCGATGAGTCCGATTCGCTACTATTCGCGATTCAGCATGCCGCGGTGATTCGTT 1320
Oy 1331 cccgagaccttgaacacacccgctgacacacacacacacacacacacacacacacacacac 1390
Db 1321 CCCGACACTTGTATACACACCGCCGCTCACACACACACACACACACACACACACACACAC 1380
Oy 1391 ttgttaacgtaaggagacagactactaagtgaactc 1429
Db 1381 TTGCCAAACCGCAAGACAGACTACTAAGGTGAACCTC 1419

RESULT 13

LSU12672 1494 bp DNA BCT 18-FEB-1995
LOCUS LSU12672 1494 bp DNA BCT 18-FEB-1995
DEFINITION Leptospira santarosai LT81 16S rRNA gene, partial sequence.
ACCESSION U12672
VERSION U12672.1 GI:558931
KEYWORDS

SOURCE Leptospira santarosai.
ORGANISM Leptospira santarosai.
Bacteria; Spirochaetales; Leptospiraceae; Leptospira.

REFERENCE 1 (bases 1 to 1494)
Ralph, D. and McClelland, M.

Phylogenetic evidence for horizontal transfer of an intervening
sequence between species in a Spirochaete genus

J. Bacteriol. 176 (19), 5962-5987 (1994)

95014031

2 (bases 1 to 1494)

McClelland, M.

Direct Submission

Submitted (25-JUL-1994) Michael McClelland, California Institute of
Biological Research, 11109 North Torrey Pines Road, La Jolla, CA
92037, USA

FEATURES

source Location/Qualifiers
1..1494

/organism="Leptospira santarosai"

/strain="serogroup Tarassovi, subserogroup Atlantae,
serovar atlantae, strain LT81"

/isolate="SA39"

/db_xref="taxon:28183"

<1..1494

/product="16S rRNA"

BASE COUNT 383 a 332 c 453 g 311 t 15 others

ORIGIN

Query Match 83.9% Score 1239.8; DB 1; Length 1494;
Best Local Similarity 93.2%; Pred No. 8.8e-12;
Matches 1363; Conservative 0; Mismatches 92; Indels 7; Gaps 7;

Oy 1 gatcatgctcgaactaaacgctcgtcgcgctccttaacaatcgagcgggtag 60
Db 22 GATCCTGCTCAAGCACTACGCTGCGCGCGCTTAAACATGCAAGCGAGTAG 81
Oy 61 caatactgagcgagcgaagcggtgagtaacacgt-ggtaactctcccgagcttgagata 119
Db 82 CAATACTCAAGCGGAGGAGGAGTAAACACGTTGATTAATCTTCCTTCGAGTCTGGGATA 141
Oy 120 acttcgcaagaagaaactaataacgagatagtccttgtagtacacaagatgatagata 179
Db 142 ACTTTCGAAAGGAGGAGGAGTAAATCTGATATGTCGATATGATGATGATGATGATGATG 201
Oy 180 aagattatctgtagatgagacccgagcgagtagtagttagtgaagaaatgagtc 239
|||||

Db 202 AAGATTGATGCTCGAGATGAGACCCCGCTCCGATTAGCTAGTTGGTGAAGGCTC 261
Oy 240 accaagcgacagatcggtgagcgcgctgagaggtgctccgacacaaatgaaatgagaca 299
Db 262 ACCAAGGCGACGATCGGTGAGCGGCTGAGAGGAGTGTGGCCACATGAACTGAGACA 321
Oy 300 cgttcatactcctcgaaggaggaagcagcttaagaactctgctcaatctgaggggaacctga 359
Db 322 CGGTCACTACT-CTACGGGAGGCGAGCAGTTAAGAACTTGTCTCAATGGGGAGACCTTA 380
Oy 360 agcagcgacgcgcgtgaaacgaagaagtcctcgatctgtaaatgtaactatgagca-ggaa 418
Db 381 AGCAGCGACGCGCGCGTGAAGATGAAGGCTCTTCGATTTTAAAGTTCAATTAAGCAGGAA 440
Oy 419 aataagcagcaatgtagatgtagtgccttcgctcaagaacccgctcaactacgttcagaca 478
Db 441 AATTAAGACGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 500
Oy 479 gccgcgtaatacgtatggtgcaagcgtctgctcgatcatcttgagcgtaaaaggtgctga 538
Db 501 GCGCGGTATATAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 560
Oy 539 ggcgagcttgtaagtcaggtgtaaaaactcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 598
Db 561 GCGGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 619
Oy 599 tacaagctgaggttggtgagagagcagtggaattccaggtgaggtgaggtgaggtgaggt 658
Db 620 TATGTCCTGGAATTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 679
Oy 659 atalcggaagcaac 718
Db 680 ATATCTGAGAGAC 739
Oy 719 cgaagcggtggtgagtaaaacgaggaatagatacccggtatccacgcgcgcgcgcgcgcgc 778
Db 740 CGAAGCGTGGGAGTGAAGCAAGGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 799
Oy 779 taccagcttggtgaggttggttaaacctcagtaaacacacacacacacacacacacacac 838
Db 800 TACCACTGTTGGGGGTTTTTAACCTCACTGATACGAA-CTAACGAGATTAAATGATGACGCT 858
Oy 839 gggagctatgctcgaagagtgaaactcaagaagatgacgggggtcgcgaacgagcgtg 898
Db 859 GGGAGCTATGCTCGCAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 918
Oy 899 gagcagctggttaactgagatgatacccaaaaacacacacacacacacacacacacac 958
Db 919 GACCATGTTGTTTAAATTCGATGATGATGATGATGATGATGATGATGATGATGATGATG 978
Oy 959 aatcagtgagagatataagcttcggcgagattcaaggtgctgacatggtgtgtgta 1018
Db 979 AATCATGATGAGATACATGAG-CTTCGGGCGCTTACAGAGTGTCTCATGTTGTGTCTCA 1037
Oy 1019 gctcgtgctgagatggttggttaagtcgcgcgaagcgaagcgaacccctacgtatgtg 1078
Db 1038 GCTCGGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1097
Oy 1079 cta-ccttaagttgggcaactgtaacgaactcgcggtgacaaacgcgaggaagcgagga 1137
Db 1098 CCATCAATTTAGTTGGGCACTCGTAAGAACTCGCGGTGACAAACCCGAGAGAAAGCGGGGA 1157
Oy 1138 tgaactcaaaactcctatggtcctatgctcgaagggcgaacagctgcttaaatgagcgagata 1197
Db 1158 TGACGTTNN 1217
Oy 1198 cagaagctcgcgaactcgcgaagagagcgaactcctaaagtcggtgccagttcgagat 1257
Db 1218 CAAAGGATGAGCAACATCGGAGGAGGAGGAGTAAATCTCAAAAAGCGGATCCAGTTCCGATT 1277
Oy 1258 ggggtctgcaactcgaccccaataagtcggaatcgcttagtaacgagatcagcagcg 1317
Db 1278 GAGTCTGCACATCGACTCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1337


```
RESULT 15
L116SRN      1396 bp      mRNA      BCT      24-AUG-1993
LOCUS        L116SRN
DEFINITION   L116SRN
ACCESSION    Z12817
VERSION      Z12817.1 GI:44003
KEYWORDS     16S ribosomal RNA; ribosomal RNA.
SOURCE       Leptospira interrogans.
ORGANISM     Bacteria; Spirochaetales; Leptospiraceae; Leptospira.
REFERENCE    1 (bases 1 to 1396)
AUTHORS      Bryden, J., Gatehouse, L., Gatehouse, J. A. and Hooke, J. V.
TITLE        Automated PCR cycle sequencing of 16S ribosomal RNA genes
JOURNAL      Unpublished
REFERENCE    2 (bases 1 to 1396)
AUTHORS      Hooke, J. V.
TITLE        Direct Submission
JOURNAL      Submitted (19-JUN-1992) Hooke, J. V., Public Health Laboratory
              Service, Leptospira Reference Laboratory, Stonebow Road, Hereford,
              Herefordshire, United Kingdom, HR1 2ER
              Location/Qualifiers
FEATURES     source
              1..1396
              /organism="Leptospira interrogans"
              /strain="leptohaemorrhagiae RGA"
              /db_xref="taxon:173"
              rRNA
              1..1396
              /standard_name="16S ribosomal RNA"
BASE COUNT   365 a      319 c      418 g      294 t
ORIGIN
```

Query Match 83.5%; Score 1232.8; DB 1; Length 1396;
Best Local Similarity 94.6%; Pred. No. 1.2e-11;
Matches 1309; Conservative 0; Mismatches 72; Indels 3; Gaps 3;

```
49  cgaagcgaggtagaatactagcgcgacgagcgagtgagtaacagt-ggtaatctctcc 107
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
13  CAAGCGAGTAGCAATACACGCGCGCAACGCGGTAGTAACACGTGGTAATCTTCTCT 72
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
108 gggctcggaatactcttcgaaagaaagaaataaccgatactgctgtgtgatacaag 167
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
73  GAGTCGGGATTAATCTTCGAAAGGAGAGCTAATTAATGATGATGATGATGATGATG 132
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
168 atttgaatgaataattatctgctgagatbagcccgcgagcttagctagttgtg 227
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
133 ATTTTGGGTAAGATTATTGCTCGGAGATGAGCCCGCTCGATTAGCTAGTTGGTG 192
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
228 aggtlaatggtcaacaagcgacgacgtcgtaacgagcgcttgagaggggtgtccgcaaat 287
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
193 AGGTAAAGGCTCACCAAGCGACGATCGGTAGCGCGCTGAGAGGGGTGTCCGCCACAAT 252
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
288 ggaactggaacacggtccactcctcgaaggaagcgagcgagcttaagaatctgtcaatg 347
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
253 GGAAGTGAACACGGTCTACTCTACCGGAGGACAGCTTAAGAAATCTTGCTCAATGG 312
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
348 gggaaacacctgaagcagcgacgcgctgaaacgaagaagctcttcgaltglaaattca 407
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
313 GGGGAACCTCGAAGACGACGCGCGGTAGACGATGAAGTCTTCGGATTGTAAAGTTCA 372
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
408 ttaagca-ggaaataaagcaatgtgatagtgtaacctgtcctaagaacacggttaac 466
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
373 GTAAGCAGGGAATAAACAGCAATGTGATGTGATGATGATGATGATGATGATGATGATG 432
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
467 taagtccagcagcgagcttaagctatgtagtaagcggttcttcggaatcatctggcgt 526
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
433 TACGTCACAGCACCGCGGTAATACGTATGCTCAAGCGTTGTCGAATCATTTGGCGT 492
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
527 aaaggggtgcgtgaagcgagattgtgaatgcaggtgtgaaactcggggtcacaacccgtg 586
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
493 AAGGGTGGGTAGCGACATGTATGTAGGTGTAAGAACTCGGGGCTCAACTCGCAGCC 552
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
587 tgcacttgaactacaaagctcgtgagtttgaggagagcaagtggaatccaggtgtgagcg 646
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
Db 553 TGCATTGAACATATGTGTCGTGAGATTGGAGAGGCAAGTGGAATTCAGAGTGTACGG 612
Qy 647 tgaatgcgtagaatctctggaagaaacacagctgaggaagcgactgtcgtgtcaaac 706
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 613 TGAATGCGTAGATATCTGAGAGACACACAGGCGAAGCTTGCTGGCTTAATACT 672
Qy 707 gacgctgaagcagaaagcgttggtgtaagaaatagatacccggttaatccagcgc 766
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 673 GACGCTGAGGACAGAAAGCGTGAGTGAACGGGATTAGATACCCGGTAAATCCAGCGC 732
Qy 767 ctaaacgtgtctaacagttgttgagggttttaacctgaagtaacgaactaagata 826
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 733 CTAAAGCTTCTACACAGTTGTTGGGGGTTTAACTCAGTAACCACTAAGCGATTAA 792
Qy 827 agtagacgcctgaggacatgctcgaagagagaaacccaagaagatgaagcgaggttc 886
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 793 AGTAGACCGCTGGGAGCTATGCTCCAAAGAGAACTCAAGAAATGACGGGGGTCC 852
Qy 887 gcaacgagctgagagcatgtgttlaattcgaatgatacccaaaaacccacactgagct 946
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 853 GCACAAGCGGTGAGCATGTGGTTTAATTCGATGATACGGGAATAAACTCAGCTAGGCTT 912
Qy 947 gacatgacatgaaatcaatgtagaatalatgagccttcggcgagatccaaggtgtgca 1006
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 913 GACATGAGATGGAATCATGTAGATACATGAGCCCTTCGGGCGCTTCAAGTCTGCA 972
Qy 1007 tggtttcgtcagctcgtgtcgtgagatgttggttaagttcccgcaagcgcaacccc 1066
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 973 TGGTTGTGTCAGCTGCTGTCGTGAGATGTTGGTTAAGTCCGCAACGAGCCAACTCT 1032
Qy 1067 tatcgtatgtgcta-cctlaagtttggtcactgtgtaacgaactgcggtgacaacacgga 1125
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1033 CACCTTATGTGTCATCATTTCACTTGGGCACTGTGAAGAACTCGCGGTGACAAACGGGA 1092
Qy 1126 ggaagcgaggatgacgtcaaatccatcgaatgacctatgtccagggcacaacagtgcta 1185
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1093 GGAAGCGGGGATGAGCTAAATCTCATGCGCTTATGTAGGGCAACACACGTGCTTA 1152
Qy 1186 caatgacgaatacagaaggttcgcgaactcgaagagagagctaatcttaaaagtcgttc 1245
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1153 CAATGGCCGGTCAAAAGGATGAGCCAACTGCGAGGGGAGCTAATCTCAAAAATCCGGTTC 1212
Qy 1246 ccagttcgaattgggtctcgaactcgaaccccaatgaagtcgaatcgtagtaatcgcg 1305
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1213 CCAGTTTCGATTGAGATCTGCAACTGACATCCATGAATCGGAATCGTAATCGCGG 1272
Qy 1306 atcaagcatgcgcgtgtgaataagttcccgagacctgttaacaacgcgcgtcaacaacct 1365
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1273 ATCAGCATGCGCGGTGGAATACGTTCCGAGACCTTGTACACACCGCCGTCAACACGCT 1332
Qy 1366 gagtgaggagacccggaagtggtcttgttaaccgttaagagagacgaactaagtgaa 1425
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1333 GAGTGGGAGACACCCGAAGTGTCTTTGCCAACCCGAAGAGACAAACTAAGGTGAA 1392
Qy 1426 actc 1429
    |||||
Db 1393 ACTC 1396
```

Search completed: December 15, 2001, 03:19:43
Job time: 7673 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 01:18:20 : Search time 401.91 Seconds
(without alignments)
3150.629 Million cell updates/sec

Title: US-09-380-826A-1

Perfect score: 1477

Sequence: 1 gacatgycgcagactaac.....ccgtaatgcattcctgcag 1477

Scoring table:

IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N.Geneseq-1101:*

- 1: /SIDS2/gcgcdata/geneseq/geneseq/NA1980.DAT:*
- 2: /SIDS2/gcgcdata/geneseq/geneseq/NA1981.DAT:*
- 3: /SIDS2/gcgcdata/geneseq/geneseq/NA1982.DAT:*
- 4: /SIDS2/gcgcdata/geneseq/geneseq/NA1983.DAT:*
- 5: /SIDS2/gcgcdata/geneseq/geneseq/NA1984.DAT:*
- 6: /SIDS2/gcgcdata/geneseq/geneseq/NA1985.DAT:*
- 7: /SIDS2/gcgcdata/geneseq/geneseq/NA1986.DAT:*
- 8: /SIDS2/gcgcdata/geneseq/geneseq/NA1987.DAT:*
- 9: /SIDS2/gcgcdata/geneseq/geneseq/NA1988.DAT:*
- 10: /SIDS2/gcgcdata/geneseq/geneseq/NA1989.DAT:*
- 11: /SIDS2/gcgcdata/geneseq/geneseq/NA1990.DAT:*
- 12: /SIDS2/gcgcdata/geneseq/geneseq/NA1991.DAT:*
- 13: /SIDS2/gcgcdata/geneseq/geneseq/NA1992.DAT:*
- 14: /SIDS2/gcgcdata/geneseq/geneseq/NA1993.DAT:*
- 15: /SIDS2/gcgcdata/geneseq/geneseq/NA1994.DAT:*
- 16: /SIDS2/gcgcdata/geneseq/geneseq/NA1995.DAT:*
- 17: /SIDS2/gcgcdata/geneseq/geneseq/NA1996.DAT:*
- 18: /SIDS2/gcgcdata/geneseq/geneseq/NA1997.DAT:*
- 19: /SIDS2/gcgcdata/geneseq/geneseq/NA1998.DAT:*
- 20: /SIDS2/gcgcdata/geneseq/geneseq/NA1999.DAT:*
- 21: /SIDS2/gcgcdata/geneseq/geneseq/NA2000.DAT:*
- 22: /SIDS2/gcgcdata/geneseq/geneseq/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1477	100.0	1477	19	AAV58896	L. fainei nucleotide
2	827	56.0	1532	22	AAV58896	Cycloclasticus pug
3	824.8	55.8	1528	22	AAV58896	Cycloclasticus pug
4	815.6	55.2	1529	22	AAV58896	Cycloclasticus pug
5	815	55.2	1533	22	AAV58896	Bacillus subtilis
6	814.6	55.2	1535	20	AAV58896	16S rRNA gene frag
7	814.4	55.1	1485	22	AAV58896	16S rRNA, Propel.
8	813.4	55.1	1512	20	AAV58896	B. subtilis partia
9	811.6	54.9	1529	20	AAV58896	16S rRNA gene frag
10	810.2	54.9	1508	20	AAV58896	Bacillus pectate I
11	810.2	54.9	1508	21	AAV58896	Nucleotide sequenc

12	810	54.8	1529	20	AAV58896	16S rDNA gene frag
13	808.4	54.7	1529	20	AAV58896	16S rDNA gene frag
14	808.4	54.7	1529	20	AAV58896	16S rDNA gene frag
15	807.8	54.7	1513	20	AAV58896	B. cereus BCM 4 pa
16	805.4	54.5	1521	16	AAV58896	16S ribosomal DNA
17	805.4	54.5	1521	16	AAV58896	P. cepacia 16S rRNA
18	805.2	54.5	1460	22	AAV58896	16S rDNA, SBR220.
19	805	54.5	1528	20	AAV58896	16S rDNA gene frag
20	805	54.5	1535	20	AAV58896	16S rDNA gene frag
21	803.8	54.4	1536	17	AAV58896	Pseudomonas testos
22	802.8	54.4	1477	20	AAV58896	Eubacterium sp. 16S
23	802.4	54.3	1627	22	AAV58896	Bacillus subtilis
24	801.8	54.3	1512	20	AAV58896	B. cereus almost c
25	801.6	54.3	1555	17	AAV58896	rRNA gene (5tp. aur
26	801.6	54.3	1555	19	AAV58896	Staphylococcus aur
27	801.4	54.3	1450	21	AAV58896	Nucleotide sequenc
28	801.2	54.2	1526	22	AAV58896	Cycloclasticus pug
29	800.8	54.2	3169	22	AAV58896	S. epidermidis gen
30	800.8	54.2	3308	22	AAV58896	S. epidermidis gen
31	800.8	54.2	3657	22	AAV58896	S. epidermidis gen
32	800.4	54.2	1532	13	AAV58896	16S rRNA gene from
33	798	54.0	1460	22	AAV58896	R. tenuis rDNA #1.
34	797.8	54.0	1460	22	AAV58896	R. tenuis rDNA #2.
35	797.6	54.0	3821	22	AAV58896	S. epidermidis gen
36	796.4	53.9	1506	20	AAV58896	Bacillus species P
37	795.6	53.9	1532	15	AAV58896	16S rRNA gene. Bo
38	794.8	53.8	1446	17	AAV58896	Root Stimulating B
39	794.8	53.8	1535	19	AAV58896	Burkholderia cepac
40	793.8	53.7	1459	22	AAV58896	16S rDNA, SBRB34.
41	793.6	53.7	1535	20	AAV58896	Nitrospira 16S rDN
42	792	53.6	349980	22	AAV58896	C glutamicum codin
43	792	53.6	349980	22	AAV58896	C glutamicum codin
44	792	53.6	349980	22	AAV58896	C glutamicum codin
45	792	53.6	349980	22	AAV58896	C glutamicum codin

ALIGNMENTS

RESULT 1	
ID	AAV58896
AAV58896	standard; DNA; 1477 BP.
AC	AAV58896;
XX	
DT	20-JAN-1999 (first entry)
XX	
DE	L. fainei nucleotide sequence.
XX	
KW	Infection: pathogenic Leptospira; protective immunity; therapy;
KW	diagnosis; ss.
XX	
OS	Leptospira fainei.
XX	
PN	W09840099-A1.
XX	
PD	17-SEP-1998.
XX	
PF	06-MAR-1998; 98WO-AU00145.
XX	
PR	07-MAR-1997; 97AU-0005494.
XX	
PA	(AGRI-) AGRIC VICTORIA SERVICES PVT LTD.
PA	(PIGR-) PIG RES & DEV CORP.
XX	
PI	Chappel RJ;
XX	
DR	WPI; 1998-520791/44.
XX	
PT	New isolated pathogenic Leptospira bacterium - useful for, e.g
PT	developing products for conferring protective immunity, and for
XX	prophylactic or therapeutic treatment
XX	

PS Claim 15; Page 69-70; 94pp; English.
 CC This sequence represents a Leptospira DNA sequence isolated from the
 CC pathogenic Leptospira (LS) bacterium of the invention. The bacterium
 CC belongs to serogroup Hurstbridge or serovar hurstbridge or the species
 CC L. falnei. The LS bacterium can be used for conferring protective
 CC immunity against pathogenic LS bacteria in humans or animals. The
 CC bacteria can also be used for prophylactic or therapeutic treatment of LS
 CC infections. The DNAs and antibodies may also be used for detection and
 CC diagnosis of past or present LS infection.

XX Sequence 1477 BP; 390 A; 334 C; 439 G; 314 T; 0 other;

Query Match 100.0%; Score 1477; DB 19; Length 1477;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1477; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gataatggtcgaactaagcgtggtggtggtttaaacaatgaaatgcaggggtag 60
 DB 1 gataatggtcgaactaagcgtggtggtggtttaaacaatgaaatgcaggggtag 60
 QY 61 caataaccagcgcaagcggttgagtaacacgttgtaattctcccgagtcggtataa 120
 DB 61 caataaccagcgcaagcggttgagtaacacgttgtaattctcccgagtcggtataa 120
 QY 121 ctcccgaaagaaagcctaataccggaatgctcgttgatcacaaagattgataagtaa 180
 DB 121 ctcccgaaagaaagcctaataccggaatgctcgttgatcacaaagattgataagtaa 180
 QY 181 agattatgtcttgagatcgagcccgccgagtaagtaagtcgtgagtaagtcgtca 240
 DB 181 agattatgtcttgagatcgagcccgccgagtaagtaagtcgtgagtaagtcgtca 240
 QY 241 ccaagcgcaagcgtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 300
 DB 241 ccaagcgcaagcgtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 300
 QY 301 ggtcattatctcttaag 360
 DB 301 ggtcattatctcttaag 360
 QY 361 ggaagcgacgcgcgtgagacgaaagaggtcttcggaatgtaagtcataagcagaanaa 420
 DB 361 ggaagcgacgcgcgtgagacgaaagaggtcttcggaatgtaagtcataagcagaanaa 420
 QY 421 ataagcaagaatgtagttagttagttagttagttagttagttagttagttagttagt 480
 DB 421 ataagcaagaatgtagttagttagttagttagttagttagttagttagttagttagt 480
 QY 481 cgcggttaacgtatgtagttagttagttagttagttagttagttagttagttagttagt 540
 DB 481 cgcggttaacgtatgtagttagttagttagttagttagttagttagttagttagttagt 540
 QY 541 cggatttgaatgtagttagttagttagttagttagttagttagttagttagttagttagt 600
 DB 541 cggatttgaatgtagttagttagttagttagttagttagttagttagttagttagttagt 600
 QY 601 caagcttgaggttgag 660
 DB 601 caagcttgaggttgag 660
 QY 661 atctgaggaacaacagctgaggaagcgagctggtggtcacaacagcgtgaggaagcag 720
 DB 661 atctgaggaacaacagctgaggaagcgagctggtggtcacaacagcgtgaggaagcag 720
 QY 721 aaagcggtggttagtaacggaggttagataccggtaaccggttaaccggttaacggttctca 780
 DB 721 aaagcggtggttagtaacggaggttagataccggtaaccggttaaccggttctca 780
 QY 781 cgaagcttgaggtttagtaacgttagtaacgttagtaacgttagtaacgttagtaacgttag 840
 DB 781 cgaagcttgaggtttagtaacgttagtaacgttagtaacgttagtaacgttagtaacgttag 840

QY 841 ggactatgtctgcaagagtgaaactcaaaaggaatgacgggggtccgcacaagcgtgtga 900
 DB 841 ggactatgtctgcaagagtgaaactcaaaaggaatgacgggggtccgcacaagcgtgtga 900
 QY 901 gcatgtgttaattatgataatgatacccaaaactcaactggtggtgataatgataatgata 960
 DB 901 gcatgtgttaattatgataatgatacccaaaactcaactggtggtgataatgataatgata 960
 QY 961 tcatgtagagataatgagccttcggtgagatcacaggtgctgtagttagttagttagttagt 1020
 DB 961 tcatgtagagataatgagccttcggtgagatcacaggtgctgtagttagttagttagttagt 1020
 QY 1021 tctgtgctgtagatgt 1080
 DB 1021 tctgtgctgtagatgt 1080
 QY 1081 accttaagttggcactgtgtgacgaactgcggttgacaacacggaaggaagcggtgata 1140
 DB 1081 accttaagttggcactgtgtgacgaactgcggttgacaacacggaaggaagcggtgata 1140
 QY 1141 cgtcaaatctctatggtccttatgttcagagggccacacagctgtacatgtgacgatacag 1200
 DB 1141 cgtcaaatctctatggtccttatgttcagagggccacacagctgtacatgtgacgatacag 1200
 QY 1201 aggttcgcaactgcgaagagggaggttaattctctaaagtcgtgtccagttcggtatgg 1260
 DB 1201 aggttcgcaactgcgaagagggaggttaattctctaaagtcgtgtccagttcggtatgg 1260
 QY 1261 gtcgcgaactgcgaagagggaggttaattctctaaagtcgtgtccagttcggtatgg 1320
 DB 1261 gtcgcgaactgcgaagagggaggttaattctctaaagtcgtgtccagttcggtatgg 1320
 QY 1321 tgaatagttcccgagactgtgtacacacggttcacacacactgtgtgtgtgtgtgtgtgt 1380
 DB 1321 tgaatagttcccgagactgtgtacacacggttcacacacactgtgtgtgtgtgtgtgtgt 1380
 QY 1381 gaagtgttcttgttgaacggaagagagagagagagagagagagagagagagagagag 1440
 DB 1381 gaagtgttcttgttgaacggaagagagagagagagagagagagagagagagagagag 1440
 QY 1441 gaagtgttcttgttgaacggaagagagagagagagagagagagagagagagagagag 1477
 DB 1441 gaagtgttcttgttgaacggaagagagagagagagagagagagagagagagagagag 1477

RESULT 2
 AAF76234
 ID AAF76234 standard; DNA; 1532 BP.
 AC AAF76234;
 DT 05-JUN-2001 (first entry)
 XX
 DE Cycloclasticus puegetii 16S rRNA gene sequence, SEQ ID NO:1.
 XX
 KM 16S rRNA gene; ribosomal RNA gene; petroleum degrading bacterium;
 KM hydrocarbon; detection; quantitation; identification; oil spill;
 KM polluted environment; environmental decontamination; ds.
 OS Cycloclasticus puegetii.
 XX
 PN W0200114587-A1.
 PD 01-MAR-2001.
 XX
 PF 24-AUG-2000; 2000WO-JP05711.
 XX
 PR 25-AUG-1999; 99JP-0237818.
 XX
 PA (AGEN) AGENCY OF IND SCI & TECHNOLOGY.
 PA (NISH-) NISHIMATSU CONSTRUCTION CO LTD.
 PA (NYRL-) NYR LOGISTICS TECHNOLOGY INST CO LTD.

XX Maruyama A, Higashihara T, Ishiwata H, Fujita T;
 XX WPI: 2001-218458/22.
 DR
 XX
 PT Detection and quantitation of microorganism having specific function
 PT and its gene from natural environment, with identification of
 PT petroleum-digesting bacterium cycloclasticus for treating oil spillage
 PT
 XX
 PS Claim 11: Page 40-41: 54pp: Japanese.

XX The invention relates to a method for detecting, identifying and
 CC quantitating a microorganism which has a specific function, or DNA
 CC therefrom, from the environment. The method comprises broadly classifying
 CC the microorganism after serial dilution; amplifying and cloning specific
 CC gene domains using extracted DNA as a template; examining the differences
 CC between the gene domains cloned and determining their base sequences; and
 CC positively identifying the microorganism from the base sequence data.
 CC The invention also relates to a method to evaluate the function of such
 CC an organism in various of environments, especially polluted
 CC environments (particularly those contaminated with oil or harmful
 CC chemicals) using migratory analysis. The methods enable the detection
 CC and quantitation of a microorganism obtained from a polluted site in
 CC order to identify strains capable of degrading petroleum and harmful
 CC chemicals. Such organisms are particularly used for environmental
 CC decontamination. The invention is particularly concerned with the
 CC detection and quantitation of petroleum-degrading bacteria of the genus
 CC Cycloclasticus. Cycloclasticus puegetii 16S rRNA gene sequences
 CC (AA76334-AA76237) are specifically claimed, as are probes (AA76238-
 CC AA76240) which hybridise to these sequences. The present sequence
 CC is one of the claimed Cycloclasticus puegetii 16S rRNA (ribosomal
 CC RNA) gene sequences.

Sequence 1532 BP: 392 A: 352 C: 473 G: 314 T: 1 other:

Query Match 56.0%; Score 827; DB 22; Length 1532;
 Best Local Similarity 76.0%; Pred. No. 5,1e-247;
 Matches 1138: Conservative 1; Mismatches 316; Indels 42; Gaps 8;

QY 1 gatactgctcgaactaactcgtgctgctcctaacaatgacgaagtcgagcggtg 60
 DB 8 gatctgctcagatgacgtgctgctgctcctaacaatgacgaagtcgagcggtg 67
 QY 61 caatc-----ctagcgcggaacgggtggtatcaacgttgtaattctt 102
 DB 68 gatactgctcgtcagcggtgctgctgctcctaacaatgacgaagtcgagcggtg 127
 QY 103 cctccgagctcgtgataacttcgaaaggaagtaacggaatgctgctgctgac 162
 DB 128 ccgagatgagggggaacactcctcgtgaaacgtctgataacggaatcctccgcgg 185
 QY 163 acaagaattgatagtataaagaattatgcttgagatgagccgcgcgagattagat 222
 DB 186 gaaagaacggggaacctcgtgctgctcctaataatgagctgacacgggagattag 245
 QY 223 tggtagagtaactgctcaccgaagcgagatcgtagcggtcgtgaggggtgtccgcgc 282
 DB 246 tggtagagtaactgctcaccgaagcgagatcgtagcggtcgtgaggggtgtccgcgc 305
 QY 283 acaatgaaactagagacgctcactactcctacgagggagcagcagatlaagaatctgtc 342
 DB 306 acacttggagctgagacgctcaccgaagcgagatcgtagcggtcgtgaggggtgtccgcgc 365
 QY 343 aatgggggaacccctgaacgagcgcgcgtgagcgaagaagctctcgtgattaa 402
 DB 366 aatgggggaacccctgaacgagcgcgcgtgagcgaagaagctctcgtgattaa 425
 QY 403 gtccataggcaggaataaagcagcagcaatgtgattgtaacttgctta----- 452
 DB 426 gcaacttcagtaggaggaataaagcttaaggtataataaccttaggcctcgaagctta 485

QY 453 -----aagcaccgctaaactacgtgcccagcagccggttaactgctgcaacgct 506
 DB 486 cagaagaagcaccgctaaactacgtgcccagcagccggttaactgctgcaacgct 545
 QY 507 tgttcggaactatggcggttaaaagggtgctgtagcggaatttgaagtcaggtgaaac 566
 DB 546 taatcggaactatggcggttaaaagggtgctgtagcggaatttgaagtcaggtgaaac 605
 QY 567 tgcgggtcaaacggtggtcgtcgtggaactgaagaatgaggttggtggaggaagc 626
 DB 606 cccgggtcaaacggtggtcgtcgtggaactgaagaatgaggttggtggaggaagc 665
 QY 627 tgaattccaggtgtagcggtgaaatgctgataatcttgagaacacacagctggcgaa 686
 DB 666 tgaattccaggtgtagcggtgaaatgctgataatcttgagaacacacagctggcgaa 725
 QY 687 cgaattgctgctcaaaactgacgctgaggaagcagaagcgtggtgtagtaacgggattag 746
 DB 726 cgcctcctggaacacacagctgaggtgaggaagcgtggtgtagtaacgggattag 785
 QY 747 ataccggtaactccagcctaaacgctgctacaggtggttgggggttttaacctca 806
 DB 786 ataccggtaactccagcctaaacgctgctacaggtggttgggggttttaacctca 845
 QY 807 gtaacgaactaactgagattagtagacgcgctgggggactatctcgaagagtgaaact 866
 DB 846 ggtggtgca-staacgaataaagttagacgcgctgggggagtagcgcgcaaggtcaaac 904
 QY 867 aagaagaattgagcggtgctccgacaagcgtgtagagcatgtgttaattcgatgatacc 926
 DB 905 aatgaattgagcggtgctccgacaagcgtgtagagcatgtgttaattcgatgatacc 964
 QY 927 caaaacgtcactggtggtcgtgcatgcatcga-atcagtatgagatatagacctcg 985
 DB 965 gaagaaccttaactcactcgtgatacagaagaacttctagagatgattgtgacctcg 1024
 QY 986 ggcagatt--cacaggtgctgcatggtgtgctcagctcgtgctgtagagatgttggtta 1043
 DB 1025 ggaactctgatacaggtgctgcatggtgtgctcagctcgtgctgtagagatgttggtta 1084
 QY 1044 agtccgcaacagcggaacccctatcgtatgttgcctacc-ctaaattggcagctgctac 1102
 DB 1085 agtccgcaacagcggaacccctatcgtatgttgcctacc-ctaaattggcagctgctac 1144
 QY 1103 gaaactcgtgtagcaaacacgggaggaagcggtgagatgagcttaaatccctcagcttga 1162
 DB 1145 ggaactcgtgtagcaaacacgggaggaagcggtgagatgagcttaaatccctcagcttga 1204
 QY 1163 tctcagggcacaacacgctgctatacaatgtagcgaatacagaagggtcgcgaacgtag 1222
 DB 1205 tggtagggctatacacaacgctgctatacaatgtagcgaatacagaagggtcgcgaacgtag 1264
 QY 1223 gagctaatctctaaagatcgtgctccagctcgtgattgggtgtcgcgaactcgaccctatga 1282
 DB 1265 aagctaatctctctaaagatcgtgctccagctcgtgattgggtgtcgcgaactcgaccctatga 1324
 QY 1283 gtcgcagctcgtgtagtaactcgcgagatcagcagctcgcgggtgaaatagcttcgcgagctgt 1342
 DB 1325 gctcgagatcgtgtagtaactcgcgagatcagcagctcgcgggtgaaatagcttcgcgagctgt 1384
 QY 1343 acacacgcgcgtlcaacacacacgtgagtgaggagcaccggaagtgtgtctgttaacgcta 1402
 DB 1385 acacacgcgcgtlcaacacacacgtgagtgaggagcaccggaagtgtgtctgttaacgcta 1443
 QY 1403 aggaagcaactactaagttagaactcgtlaaagggggtgaagtctgtaaacgaagctac 1459
 DB 1444 ggaagcgcgtcaacacacttggatcactgactgaggtggaagtcgtaaacgaagctac 1500

RESULT 3
 AAF76235
 ID AAF76235 standard; DNA: 1528 BP.
 XX

AA76235;
 05-JUN-2001 (first entry)
 Cycloclasticus puegell 16S rRNA gene sequence, SEQ ID NO:2.
 16S rRNA gene; ribosomal RNA gene; petroleum degrading bacterium;
 hydrocarbon; detection; quantitation; identification; oil spill;
 polluted environment; environmental decontamination; ds.
 Cycloclasticus puegell.
 W0200114587-A1.
 01-MAR-2001.
 24-AUG-2000: 2000WO-JP05711.
 25-AUG-1999: 99JP-0237818.
 (AGEN) AGENCY OF IND SCI & TECHNOLOGY.
 (NISH-) NISHIMATSU CONSTRUCTION CO LTD.
 (NYKL-) NYK LOGISTICS TECHNOLOGY INST CO LTD.
 Maruyama A, Higashihara T, Ishiwata H, Fujita T;
 WPI: 2001-218458/22.
 Detection and quantitation of microorganism having specific function
 and its gene from natural environment, with identification of
 petroleum-digesting bacterium cycloclasticus for treating oil spillage
 Claim 11: Page 41-42; 54pp: Japanese.
 The invention relates to a method for detecting, identifying and
 quantitating a microorganism which has a specific function, or DNA
 therefrom, from the environment. The method comprises broadly classifying
 the microorganism after serial dilution; amplifying and cloning specific
 gene domains using extracted DNA as a template; examining the differences
 between the gene domains cloned and determining their base sequences; and
 positively identifying the microorganism from the base sequence data.
 The invention also relates to a method to evaluate the function of such
 an organism in various of environments, especially polluted
 environments (particularly those contaminated with oil or harmful
 chemicals) using migratory analysis. The methods enable the detection
 and quantitation of a microorganism obtained from a polluted site in
 order to identify strains capable of degrading petroleum and harmful
 chemicals. Such organisms are particularly used for environmental
 decontamination. The invention is particularly concerned with the
 detection and quantitation of petroleum-degrading bacteria of the genus
 Cycloclasticus. Cycloclasticus puegell 16S rRNA gene sequences
 (AA76234-AA76237) are specifically claimed, as are probes (AA76238-
 AA76240) which hybridise to these sequences. The present sequence
 is one of the claimed Cycloclasticus puegell 16S rRNA (ribosomal
 RNA) gene sequences.
 Sequence 1528 BP; 394 A; 343 C; 467 G; 323 T; 1 other:
 Query Match 55.8%; Score 824.8; DB 22; Length 1528;
 Best Local Similarity 76.0%; Pred. No. 2.4e-246;
 Matches 1135; Conservative 1; Mismatches 318; Indels 40; Gaps 8;

128 taatagtgtggaacaactcgtgtgaataaaccaggtactatccgataatccctacgggca 187
 165 aagatttgataggtaaaagattatgtcttgagatagcccgccgattagctgtg 224
 188 agca--gggacacctggtgctgtcgtcttaataagatagctatgtcgtatagctagt 245
 225 gtaggtgaatggtctcaacgaagcgaacgtgtagccggtcctagaaggtgtccgccc 284
 246 gtaggtgaatggtctcaacgaagcgaacgtgtagccggtcctagaaggtgtccgccc 305
 285 aatggaacttgagaacacgtgtccatactctacggaagcagcttaagaactctgtctca 344
 306 actggaacttgagaacacgtgtccatactctacggaagcagcttgggaatatgtgacaa 365
 345 tgggggaacccctgaagcagcgcgcgttgaaacgaagaagctcttcgttaagt 404
 366 tggaggaactcgtatgcacaaatgcgcgtgtgtgaagaagccttaagggtgttaagc 425
 405 tcatlaagcaggaataaagcagcaatgtgatgtgtacgtgtccta----- 452
 426 acttcagtagggaggaataagtttaagtttaacttagccctgaagcttactacag 485
 453 ---aagcaccgctaaactacgttccagcagccgcgttaactatgtgtcaagcgttgt 509
 486 aagaagcaccgctaaactacgttccagcagccgcgttaactatgtgtcaagcgtttaa 545
 510 tccgaatcatttggcgttaaaaggtgtgcttagcgaattgttaagcaggtgtgaaacgtc 569
 546 tccgaatcatttggcgttaaaaggtgtgcttagcgaattgttaagcaggtgtgaaacgtc 605
 570 gggctcaaccgtgtgctgtcacttgaaactaacaagtctgtgaatttggagaagcagttg 629
 606 gggctcaaccgtgtgctgtcacttgaaactaacaagtctgtgaatttggagaagcagttg 665
 630 aatccaggtgtgagcgttgaatgtcgtatgtatcttgagaagaaacaggtgccaagcga 689
 666 aattcgggtgtgagcgttgaatgtcgtatgtatcttgagaagaaacaggtgccaagcga 725
 690 ctctgtgctcaaaacgtgacgttgaagcagcaagcgtgtggttagtaaacgggtatagata 749
 726 ctctgtgctcaaaacgtgacgttgaagcagcaagcgtgtggttagtaaacgggtatagata 785
 750 ccccggtatccacgcgcctaaacgttgtctaccagttgttgggggttttaacctca 809
 786 ccccggtatccacgcgcctaaacgttgtctaccagttgttgggggttttaacctca 845
 810 acgaacctaaacgttgaatgaatgaacgcgcctgtggaactgtctccgcaagaagttaacctaaa 869
 846 gtgca-staaacgttgaatgaatgaacgcgcctgtggaactgtctccgcaagaagttaacctaaa 904
 870 ggaattgacgggtgtccgcacaacgtgtgagcagtggttttaatttcgatatacccaa 929
 905 tgaattgacgggtgtccgcacaacgtgtgagcagtggttttaatttcgatatacccaa 964
 930 aaacctcaacctgtgacatgtgactcga-atcaatgaagatatatgaaccttgggc 988
 965 gaaccttacccttgaacatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 1024
 989 agatt--cacaggtgtgcatgtgtgtgtcagctcgtgtcgtatgaatgttgggttaagt 1046
 1025 actcgtatagcaggt 1084
 1047 cccgcaacgaacgaacccctatcgtatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1105
 1085 cccgtaacgaacgaacccctatcgtatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1144
 1106 actgctgtgtgacaacacgtgaggaagcgtggtgatatgacgttcaaatctctatgtgtgtgtgt 1165
 1145 actgctgtgtgacaacacgtgaggaagcgtggtgatatgacgttcaaatctctatgtgtgtgtgt 1204
 1166 ccagggtcacaacacgt 1225

DB 1205 gtagggctacacacgtgtctacaaatgycggtacagagggcgcaaacctcgagagtaag 1264
OY 1226 ctaattctctaaagtcggtcccaagtcggtatgggtctgcgaactcgaccocatgaagtc 1285
DB 1265 ctaattctctaaagtcggtcccaagtcggtatgggtctgcgaactcgaccocatgaagtc 1324
OY 1286 ggaatcgtcagtaaatcgagatcagatccgcgcgtggaatcgtccgcgacttataa 1345
DB 1325 ggaatcgtcagtaaatcgagatcagatccgcgcgtggaatcgtccgcgacttataa 1384
OY 1346 caccgcgcgtac 1405
DB 1385 caccgcgcgtac 1442
OY 1406 agac 1459
DB 1443 agcgcgtac 1496

RESULT 4

AAF76236
ID AAF76236 standard; DNA; 1529 BP.

AC AAF76236;

DT 05-JUN-2001 (first entry)

DE Cycloclasticus pugettii 16S rRNA gene sequence, SEQ ID NO:3.

KW 16S rRNA gene; ribosomal RNA gene; petroleum degrading bacterium;
hydrocarbon; detection; quantitation; identification; oil spill;
KM polluted environment; environmental decontamination; ds.

OS Cycloclasticus pugettii.

PN WO200114587-A1.

PD 01-MAR-2001.

PF 24-AUG-2000; 2000WO-JP05711.

PR 25-AUG-1999; 99JP-0237818.

PA (AGEN) AGENCY OF IND SCI & TECHNOLOGY.
PA (NISH-) NISHIMATSU CONSTRUCTION CO LTD.
PA (NYKL-) NYK LOGISTICS TECHNOLOGY INST CO LTD.

PI Maruyama A, Higashihara T, Ishiwata H, Fujita T;

DR WPI; 2001-218458/22.

PT detection and quantitation of microorganism having specific function
PT and its gene from natural environment, with identification of
PT petroleum-digesting bacterium cycloclasticus for treating oil spillage

XX Claim 11: Page 42-43; 54pp; Japanese.

XX The invention relates to a method for detecting, identifying and
XX quantitating a microorganism which has a specific function, or DNA
XX therefrom, from the environment. The method comprises broadly classifying
XX the microorganism after serial dilution; amplifying and cloning specific
XX gene domains using extracted DNA as a template; examining the differences
XX between the gene domains cloned and determining their base sequences; and
XX positively identifying the microorganism from the base sequence data.
XX The invention also relates to a method to evaluate the function of such
XX an organism in various of environments, especially polluted
XX environments (particularly those contaminated with oil or harmful
XX chemicals) using migratory analysis. The methods enable the detection
XX and quantitation of a microorganism obtained from a polluted site in
XX order to identify strains capable of degrading petroleum and harmful
XX chemicals. Such organisms are particularly used for environmental
XX decontamination. The invention is particularly concerned with the

CC detection and quantitation of petroleum-degrading bacteria of the genus
CC Cycloclasticus. Cycloclasticus pugettii 16S rRNA gene sequences
CC (AAF76234-AAF76237) are specifically claimed, as are probes (AAF76238-
CC AAF76240) which hybridise to these sequences. The present sequence
CC is one of the claimed Cycloclasticus pugettii 16S rRNA (ribosomal
CC RNA) gene sequences.

XX Sequence 1529 BP; 394 A; 343 C; 466 G; 320 T; 6 other;

Query Match 55.2%; Score 815.6; DB 22; Length 1529;
Best Local Similarity 75.4%; Pred. No. 1.8e-243;
Matches 1126; Conservative 4; Mismatches 325; Indels 39; Gaps 8;

OY 1 gatcaggtctcgaactaagcgtcggtggtgttcttaaacatgcaagtcgagcggtag 60
DB 8 gatcaggtctcgaactaagcgtcggtggtgttcttaaacatgcaagtcgagcggtag 67
OY 61 caatac-----ctagcgcgaacggtgtagtaacacgtgtatctcc 104
DB 68 atgctagctctgtagcaggtcgtagtggtgcgcggtgtagtaatgcataggaattacc 127
OY 105 tccgagtcgtggaataacttcgaaagaactaaccgtagtctctgtgatacc 164
DB 128 taacagctggggacacacctgtgtgaaaccagcttaacgcataatccctaacgggcaa 187
OY 165 aagattgtagtgaagatttctgttggaatgaagccgcgcgtttagctagtg 224
DB 188 agca---ggggaacctcggtggtctgtagtgaatgaagccgcgcgtttagctagtg 245
OY 225 gtagagtaatgtctcacaagcgcaagcagatcgtagtcggtcgtgagaggtgttcgcgcac 284
DB 246 gtgaggtgaatgtcgcac 305
OY 285 aatgagactgagacacggtctcactcctcctcaggaagcgacagtaagaattctctcaa 344
DB 306 actgagactgagacacggtcctcactcctcaggaagcgacagtaagaattctctcaa 365
OY 345 tgggggaacccctggaagcgacggtcgtgaggaagaaggtcttcgattgtaagt 404
DB 366 tggaggaacacctcgaatgcac 425
OY 405 tcataggcagaaaataagacagcaatgtgatagtatgcctcctc----- 452
DB 426 acttctagtaggaggaagaaggttaaggttaataacttaggcctcagcgttaactacag 485
OY 453 ----aagcaccggtactacgtgcacagcgcggttaatacgtat-tgtgtcaagcgttg 508
DB 486 aagaagcaccggtactcgtgcacagcgcggttaatacgtat-tgtgtcaagcgttg 545
OY 509 ttcggaatcattgggtgaaggttcgttaggggtttgtaagtagtgtgtgaaactcg 568
DB 546 atcggaatcattgggtgaaggttcgttaggggtttgtaagtagtgtgtgaaactcg 605
OY 569 cgggctcaacccgtggtcgtcacttgaactgaacacacacacacacacacacacacacacac 628
DB 606 cgggctcaacccgtggtcgtcacttgaactgaacacacacacacacacacacacacacacac 665
OY 629 gaattcaggtgtgaggtgtaaatgtcgtatatactcgtggaacacacacacacacacacac 688
DB 666 gaattcaggtgtgaggtgtaaatgtcgtatatactcgtggaacacacacacacacacacac 725
OY 689 acttctggtcgtcaaacctgagcgtgagcgaagaacggtgtagtaaacaggtatagat 748
DB 726 gctctgtgac 785
OY 749 acccgtgaatccacgaccttaaacgtgtgtctacacagttgttgggggttttaacctcagt 808
DB 786 acccgtgaatccacgaccttaaacgtgtgtctacacagttgttgggggtttttccgttagt 845
OY 809 aagcaactaagatgaatgaacggtcgtggtgactatgctgcgaaggtgaactcaa 868
DB 846 gtgtca-staacgataagttgacgcgttgggagtaagccgcgaaggttaaacactcaa 904


```
Db 856 cccttagtgcgcgaatacgaatgaacatccgccttgaggagtagcgtgcgaagactg 915
Oy 861 aaactcaaggaatgcagcggggtccgcgaagcgttgagacatgtgttattcagatg 920
   |||||
Db 916 aaactcaaggaatgcagcggggtccgcgaagcgttgagacatgtgttattcagag 975
Oy 921 ataccacaaactcactggtggttgaatc-gaatctgaatcattgagatatacgag 979
   |||||
Db 976 caacgcgaagaacatcaccaggtcttgacatctcctgacaatccatcgatagtagaagctcc 1035
Oy 980 ccttc--ggcagatcagcagtgctgcatgtgttcgtcagcctcgttcgtgagatgtt 1037
   |||||
Db 1036 ccttcggggcagagtgagcaggttggtgcatgtgttcgtcagcctcgttcgtgagatgtt 1095
Oy 1038 gggttaagcccgcaagcagcggcgaacccct-atcgatgtgttctacttaagtgtggcc 1096
   |||||
Db 1096 gggttaagcccgcaagcagcggcgaaccccttgatcttgatgtcgaagatcagttggcc 1155
Oy 1097 tggtaagaactgcgcgtgacaaacccgaggaagcgggagatgcgttcaatcctcatg 1156
   |||||
Db 1156 tctaaagtgacgtccgcgtgacaaacccgaggaagcgttgagacgttcaatcctcatg 1215
Oy 1157 ccttaatgtccagcggcacaacacgtgtcacaatgtccgacatagaggtgcgcaactgc 1216
   |||||
Db 1216 ccttaatgtccgtggtctacacacgtgtcacaatgtcagaagaaggtgcgcaacccgc 1275
Oy 1217 aagagggagatcactctcctaaagtctgcaggttcggtggtgtcgaactgacc 1276
   |||||
Db 1276 gagggttaagccaaatccacaatcgtctcagttcagttcgacgtcgaactcgaactgacgt 1335
Oy 1277 catgaagtcgaactgcgtgacaaacccgaggaagcgtgacgtgacgttaccgtg 1336
   |||||
Db 1336 cgtgaagcgtgacgtgacgtgacaaacccgaggaagcgtgacgtgacgttaccgtg 1395
Oy 1337 ccttgtaacacacccgcgtgcacacacccgtgagtgaggagacccgaggtgcttctgta 1396
   |||||
Db 1396 ccttgtaacacacccgcgtgcacacacccgtgagtgaggagacccgaggtgcttctgta 1455
Oy 1397 accgtgaaggaagacactaactaagtgtaaacgtgtaaggggtgagtcgttaacaagt 1456
   |||||
Db 1456 ccttcttaagagcagcagcgcggaaggtggaacagatgattcggggtgagtcgttaacaagt 1515
Oy 1457 acc 1459
   ||
Db 1516 agc 1518

RESULT 6
AAK83569
ID AAK83569 standard; DNA; 1535 BP.
XX
AC AAK83569;
XX
DT 21-DEC-1999 (first entry)
XX
DE 16S rDNA gene fragment from marine bacterium, isolate K3-3.
XX
KM Monitoring; oil; contamination; sea water; detection; flagellum;
KM Gram-negative bacterium; Proteobacteria; glucose; carbon source;
KM alkane; 16S rDNA gene; ds.
XX
OS Proteobacteria.
XX
PN JPL1243967-A.
XX
PD 14-SEP-1999.
XX
PF 04-MAR-1998; 98JP-0069399.
XX
PR 04-MAR-1998; 98JP-0069399.
XX
PA (KAIY-) KAIYO BIOTECHNOLOGY KENKYUSHO KK.
XX
```

```
DR WPI; 1999-564435/48.
XX
PT Monitoring of oil contamination of sea water - where oil contamination
PT is evaluated by detection of a microbe having properties from e.g.
PT having no flagellum, being a Gram-negative bacterium, belonging to
PT Proteobacteria, gamma subdivision, etc.2
XX
PS Claim 3: Page 7-8; 15pp; Japanese.
XX
CC The invention relates to a method for monitoring oil contamination of
CC sea water by detecting, in the sea water, a microbe having the following
CC properties: (1) it has no flagellum; (2) it is a Gram-negative bacterium;
CC (3) it belongs to Proteobacteria, gamma subdivision; (4) it cannot
CC assimilate glucose as a single carbon source; and (5) it efficiently
CC assimilates at least one of 10-30C n-alkanes. This sequence represents
CC a fragment of the 16S rDNA gene from the microbe of the invention,
CC isolate K3-3.
XX
SQ Sequence 1535 BP; 375 A; 350 C; 496 G; 314 T; 0 other;
```

```
Query Match 55.2%; Score 814.6; DB 20; Length 1535;
Best Local Similarity 76.0%; Pred. No. 3,76-243;
Matches 1137; Conservative 0; Mismatches 319; Indels 41; Gaps 9;
```

```
Oy 1 gatcatgctcagaactaagcgtggtggcgccgtcttaaacatcgaagtcgagcggtag 60
   |||||
Db 9 gatccgtgctcagatggaacgtcgtgcgcgacgctcaacatcgaagtcgagcggtag 68
Oy 61 caatac-----ctagcggcgaacgggtgtagtaacacgtgtaattcc 104
   |||||
Db 69 atgtagctgtaccagcgttcgagcggcgggtgtagtaacgctgagatctgc 128
Oy 105 tccgactcgggaatacttccgaagaagtaataccgataagtcctgtgtgacac 164
   |||||
Db 129 catttggtggataacttgggaaactcaagctaaactacgataacccctacggggaa 188
Oy 165 aagatttgatgaagaatttattcttggaatagcccgcgacgttaagctgattg 224
   |||||
Db 189 agaa--ggggactctgggcctgtgcagatgataagtcgctggttagctagtgtg 246
Oy 225 gtgaagtaatggtcacaagcgcagatcgtgtagccgctgtgagagtgctccgac 284
   |||||
Db 247 gtggggtaaaagcctcaagaagcagatcgtgtagctgtgtgaaggatgatacagccac 306
Oy 285 aatgaactgagaacaggtccataactcctacgggaagcagcagtgaggaaacttgctcaa 344
   |||||
Db 307 accgggactgagacacgcgcgcgtccctacggaagcagcagtgaggaaacttgacaa 366
Oy 345 tggggaacccctgaagcagcgcgcgtggaacgaagaagctcctcgatgttaagt 404
   |||||
Db 367 tggggcaacccctgataccgcgtcgtggtgtgtaagaagccttaaggtgtgtaag 426
Oy 405 tcaatgacgagaaaaata-----agcagcaatgtgatatgtactgcct 451
   |||||
Db 427 acttccagcaggaaggaagccttcgagttaatccttggaactgagcttaccctgcag 486
Oy 452 a--aaagcagcgttaactcgtgcagcgcgcgttaactagtagtggcgaagctgt 509
   |||||
Db 487 aagaagcagcgttaactcgtgcagcgcgcgttaactagtagtggcgaagctgt 546
Oy 510 tgggaatcattggcgtaaaaggctgtaggcggaatttlaagtcaaggtgtgaaactgc 569
   |||||
Db 547 tgggaattactggcgtaaaaggcgtgtaggcggttcttgaagtaagtagtgaagccccc 606
Oy 570 gggtcaaccgttgctgcaacttgaactacaagtcctggagtttgggaggaagcgtg 629
   |||||
Db 607 gggtcaaccgttggaactcatttgaacttgcagcgttagagtgtagtgagaggtg 666
Oy 630 aattcaggtgtgagcgttaaatgctgatatcgtgaagaaacacgaagcgaagcga 689
   |||||
Db 667 aattcgggtgtagcgtgaaatgctgatacgtgaagaagaacacgaagcgtg 726
```

[illegible]

```

XX 23-DEC-1999;      99AU-0004867 .
XX
XX (CRCW-) CRC WASTE MANAGEMENT & POLLUTION CONTROL.
XX
XX Hugenholz P, Crocetti GR, Tyson GW, Blackall LL:
XX
XX MPI: 2001-408656/43.
XX
XX Novel oligonucleotide probe or primer useful for detecting
PT polynophosphate accumulating organism in a sample, comprises a sequence
PT that is unique to 16S rDNA of polynophosphate accumulating organisms -
XX
XX
XX Claim 4; Fig 3; 54pp: English.
XX
XX The sequences given in AAC86021-30 represent 16S rDNA sequences from
CC polyphosphate accumulating organisms (PAOs). Sequences which are unique
CC to these 16S rDNA sequences are used to create a probe or primer for
CC detecting the relevant organisms. The primer/probe sequences are
CC useful for detecting PAO cells in a sample, by treating cells in the
CC sample to fix cellular contents, contacting fixed cells with the
CC primer/probe which is labelled with a radiolabel, a reporter group or a
CC hapten, under conditions which allow the probe to hybridize with 16S
CC rRNA within the fixed cell, removing unhybridized probe from the fixed
CC cells, and detecting the labeled probe-RNA hybrid by fluorescence in
CC situ hybridization. The primer/probe sequences are useful for
CC identifying PAOs that are capable of biologically removing
CC phosphorus from waste water. Rapid assessment of the presence of a
CC number of PAOs in a waste water sample, can be done using the primer/
CC probe sequences. They allow quick and convenient assessment of whether
CC a sludge or waste water sample includes PAOs and allows quantitation
CC of PAO cells in samples.
XX
XX Sequence 1485 BP; 367 A; 341 C; 482 G; 295 T; 0 other:
SQ

```

```

Db 479 cggcgtaactaagctgcccagcggcgtaatacgttaagctgcccagcttaatcggaaatt 538
OY 518 attggcgtaaaaggctgctgaagcggaattgtcaagtcaagctgtgaactcggcgctcaa 577
Db 539 actggcgtaaaaggctgctgaagcggaattgtcaagtcaagctgtgaactcggcgctcaa 598
OY 578 cccggtgacctgaactgaactgaactgaactgaactgaactgaactgaactgaactgaact 637
Db 599 cctgggaatgacctgaactgaactgaactgaactgaactgaactgaactgaactgaact 658
OY 658 gtttagcgggtgaactgaactgaactgaactgaactgaactgaactgaactgaactgaact 697
Db 659 gtttagcgggtgaactgaactgaactgaactgaactgaactgaactgaactgaactgaact 718
OY 698 ctcaaaactgaactgaactgaactgaactgaactgaactgaactgaactgaactgaact 757
Db 719 gttactactgaactgaactgaactgaactgaactgaactgaactgaactgaactgaact 778
OY 758 atccacgacctgaactgaactgaactgaactgaactgaactgaactgaactgaactgaact 815
Db 779 gttcaccgacctgaactgaactgaactgaactgaactgaactgaactgaactgaactgaact 838
OY 816 ctacacggaattagtagaccgctggtggaactatgctcgaagaatgaactgaactgaact 875
Db 839 ctacacggaattagtagaccgctggtggaactatgctcgaagaatgaactgaactgaact 898
OY 876 gacgggggtgcgcacaaaggggtggaagcattgattatcattatcattatcattatcatt 935
Db 899 gacgggggtgcgcacaaaggggtggaagcattgattatcattatcattatcattatcatt 958
OY 936 caccggggtgacctgacatgacatgacatgacatgacatgacatgacatgacatgacatg 990
Db 959 tacttaacctgacatgacatgacatgacatgacatgacatgacatgacatgacatgacat 1018
OY 991 attcaacaggtgctgaactgctgctgaactgctgctgaactgctgctgaactgctgctga 1050
Db 1019 gacacacaggtgctgaactgctgctgaactgctgctgaactgctgctgaactgctgctga 1078
OY 1051 caacgagcgcaacacctgctgaactgctgctgaactgctgctgaactgctgctgaact 1109
Db 1079 caacgagcgcaacacctgctgaactgctgctgaactgctgctgaactgctgctgaact 1138
OY 1110 ccggtgacacaaacggaggaagcggtggaactgaactgaactgaactgaactgaactgaact 1169
Db 1139 ccggtgacacaaacggaggaagcggtggaactgaactgaactgaactgaactgaactgaact 1198
OY 1170 ggcacacacgtgctacaaatggcgaatagaaggggtgcgcaactcgcaagaaggagctaa 1229
Db 1199 ggccttcacacgtacaaatggcgaatagaaggggtgcgcaactcgcaagaaggagctaa 1258
OY 1230 tctctaaagtgctgctccagctgctgaactgctgctgaactgctgctgaactgctgctga 1289
Db 1259 tctctaaagtgctgctccagctgctgaactgctgctgaactgctgctgaactgctgctga 1318
OY 1290 tgcgtagtaactgcggaatcgaactgcggtggaatgaactgcggtggaatgaactgcggtgga 1349
Db 1319 tgcgtagtaactgcggaatcgaactgcggtggaatgaactgcggtggaatgaactgcggtgga 1378
OY 1350 ggcgtgacacacacgtgagtggtggaagcaccggaagtgtctctgttaaccgttaagaag 1409
Db 1379 ggcgtgacacacacgtgagtggtggaagcaccggaagtgtctctgttaaccgttaagaag 1437
OY 1410 agactactaagtgaactcgttaagaagggtggaagcgttaagaagcgttaagaagcgtta 1457
Db 1438 cgtctacacacggtggtgctgactgagtggtggaagcgttaagaagcgttaagaagcgtta 1485

```

RESULT 8
 AAX82000
 ID AAX82000 standard; DNA; 1512 BP.
 XX
 AC AAX82000;

```

XX 10-SEP-1999 (first entry)
DT B. subtilis partial 16S rRNA gene sequence.
DE Taxane; pacillitaxel; plant; bacteria; mutation; cancer; 16S rRNA; ss.
KW Bacillus subtilis.
XX WO9932651-A1.
PN 01-JUL-1999.
XX 18-DEC-1998; 98WO-CA01150.
PF 22-DEC-1997; 97US-0995960.
PR (BCMB-) BCM DEV INC.
PA Bolesnot M, Gagne M, Harvey M, Helle M, Landry N;
PI Page M;
PI WPI; 1999-418940/35.
DR Production of taxane and pacillitaxel compounds
XX Disclosure; Fig 8F; 93pp; English.
XX The invention relates to a new method for the production of taxane and
CC pacillitaxel compounds that comprises culturing bacteria isolated from a
CC plant species of Taxus or bacteria produced by mutating the isolated
CC bacteria. The methods can be used for the production of taxanes such as
CC pacillitaxel, 10-deacetylcephalomannine, 7-epitaxol, 10-deacetyl-7-epi-
CC taxol, 7-epicephalomannine, baccatin III, 7-xylosyltaxol, 7-xylosyl-
CC cephalomannine, 7-epibaccatin III, 7-xylosyltaxol, 9-acetyloxy
CC taxusin, 9-hydroxy taxusin, taxane Ia, taxane Ib, taxane Ic, or taxane
CC Id (claimed). The taxane and pacillitaxel compounds can be used for the
CC treatment of cancers. The methods can provide for the production of
CC taxanes and pacillitaxel at a concentration of 1-25 mu g/L.
XX
XX Sequence 1512 BP; 377 A; 353 C; 480 G; 302 T; 0 other;
SQ

```

Query Match 55.1%; Score 813.4; DB 20; Length 1512;
 Best Local Similarity 75.4%; Pred. No. 8,6e-243;
 Matches 1133; Conservative 0; Mismatches 326; Indels 44; Gaps 8;

```

OY 1 gatactgctcagaactaactgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 55
Db 8 gatccgtgctcagaactaactgctgctgctgctgctgctgctgctgctgctgctgctgctgct 67
OY 56 -----gtagcaatcctagcggcggaaggggtgaatgaactgctgctgctgctgctgct 105
Db 68 tgggaactgctgctcctgtagttagcgcggaaggggtgaatgaactgctgctgctgctgctgct 127
OY 106 ccgaactgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 165
Db 128 gtaagaactgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 187
OY 166 agattgataagtaag-----attatgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 218
Db 188 gttcagacataaagggtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 247
OY 219 tagtggtagagtaagtgtcctacacgaagcgaactgctgctgctgctgctgctgctgctgctgctgct 278
Db 248 tagtggtagagtaagtgtcctacacgaagcgaactgctgctgctgctgctgctgctgctgctgctgctgct 307
OY 279 ggcacacatggaactgaagcagctgctacatcctcctcgggagggcagcagctgaagtaactct 338
Db 308 ggcacacatggaactgaagcagctgctacatcctcctcgggagggcagcagctgaagtaactct 367
OY 339 gctcaatgggggaaacccctgaagcagcagcgtggaacgaagaagctcttgcgattg 398

```

```

Db      368  cccgaatgacgaagctcgaaggaaacacccgcgtgcatgtaagagttccgatacg  427
Qy      399  taaagtcatgaagcagaanaaaataagcaacaaatgtagtgaacctgct-----  451
Db      428  taaagctctgtgttagggagaacaaagtcgcttcaaatagggcgaccttgaagcgt  487
Qy      452  -----aaagcacgcgttaactacgtccagcaagccgcgttaactgtagtgcaa  502
Db      488  cctaaccagaagaacgagcttaactacgtccagcaagccgcgttaactgtagtgcaa  547
Qy      503  ggcgtgttcgaagaatctggcgtaaaagtgctgaagcgattgtaactgaagtgta  562
Db      548  ggcgtgttcgaagaatctggcgtaaaagtgctgaagcgattgtaactgtagtgta  607
Qy      563  aaactcggcgctcaaacccgtgctcctgaactgaactacaagtcgtgaagtgagag  622
Db      608  aagcccccgttcaacgaggggaggtcatgtgaacctggggaacttgatgcagaagag  667
Qy      623  caagtgaatctccagtgtagcggtgaagtcgtgagatattctggagaaacccagtg  682
Db      668  aagtggaattccacgtgtgtagcgtgtagagatgtggaaggaacacacagtgagcg  727
Qy      683  aaggcgagctgtgctcaaacacgtagcgtgaagcagcaagacgtggtgaagaa  742
Db      728  aagcgagactctctgtctgttaactgaacgtgagagcgtagggagcgaaacagaa  787
Qy      743  ttagatacccggttaactacacgcttaagctgtctacacgattggtggagttt--aa  800
Db      788  ttagatacccggttagtgcacgcttaacgaatgagtgctgaagtgtaggggttccgc  847
Qy      801  cccctcagtaacgaacctaaccgaatgaacgcgcctggggaactatctcgaagaag  860
Db      848  cccttagtgctcgacgaactaactaagaacatccgcctgggagtagcgtcgaagact  907
Qy      861  aaactcaaaagaatgtgaagcggttcgcacaaagcgtgtgaagcatgtgttaactg  920
Db      908  aaactcaaaagaatgtgaagcggttcgcacaaagcgtgtgaagcatgtgttaactg  967
Qy      921  ataccaccaaaacctcaactcgtggttgacat-ggattcgaatcatgtagatatagag  979
Db      968  caacgcgaagaacctaaccgaatgttgaacatcccttgacaacccaagaagatagag  1027
Qy      980  ccttc--gggcagattccaggtgctgcatgtgtgtcgcagctcgtgtcgtgagatgt  1037
Db      1028  ccttcgggggcagagtgacaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt  1087
Qy      1038  ggggttaagtcgccgaacagagcgcaacccct-actgtatgtgtctacttaagtgg  1096
Db      1088  ggggttaagtcgccgaacagagcgcaacccctgtactgtgtcgaacatcagttgg  1147
Qy      1097  tggtagcaaacctcgcgtgtgacaacacggaggaagcggtgaacgtcaaacctcag  1156
Db      1148  tctaaggtgtagctgcgggtgtgacaacacggaggaaggtgtggatgtagcatcat  1207
Qy      1157  ccttatgttcaggggcacacacgtgtactaaatggcgcgaatacagaggtgcgaac  1216
Db      1208  cccctatgacccgtggcgtacacacgtgtactaaatggcgcgaatacagaggtgc  1267
Qy      1217  aagaggggaactatctcttaaaagtcgtgtccaggttcgagttgggtgtcctcaac  1276
Db      1268  gagggttaagccaactcccaaaactgtgtcagttcgtgaltcgaagtcctgcaact  1327
Qy      1277  catgaagtcggaatcgtgtaactcgcgagatcagatcgtccgcgtgtgaatacgt  1336
Db      1328  cgtgaagcgtggaatcgtgtaactcgcgagatcagatcgtccgcgtgtgaatacgt  1387
Qy      1337  ccttgcatacacacgcgcgttcaacacacgtgagtggtggagcaacccgaagtg  1396
Db      1388  ccttgcatacacacgcgcgttcaacacacgtgagtggtggagcaacccgaagtg  1447
Qy      1397  accgtgaaggaagaactacttaaggtgaaactgttaaggggtggaagtcgttaac  1456

```

```

Db      1448  ccttatgagccagccgcgaaggtgtggagacagatgtgtgggtgtagatcgtlaacaa  1507
Qy      1457  acc 1459
Db      1508  aac 1510

RESULT      9
AAK83564
ID      AAK83564 standard; DNA; 1529 BP.
XX
AC      AAK83564;
XX
DT      21-DEC-1999 (first entry)
XX
DE      16S rDNA gene fragment from marine bacterium, isolate SF-T1.
XX
KW      Monitoring; oil; contamination; sea water; detection; flagellum;
KW      Gram-negative bacterium; Proteobacteria; glucose; carbon source;
KW      alkane; 16S rDNA gene; ds.
XX
OS      Proteobacteria.
XX
PN      JF11243967-A.
XX
PD      14-SEP-1999.
XX
PF      04-MAR-1998; 98JP-0069399.
XX
PR      04-MAR-1998; 98JP-0069399.
XX
PA      (KAIY-) KAIYO BIOTECHNOLOGY KENKYUSHO KK.
XX
DR      WPI; 1999-56435/48.
XX
PT      Monitoring of oil contamination of sea water - where oil contamination
PT      is evaluated by detection of a microbe having properties from e.g.
PT      having no flagellum, being a Gram-negative bacterium, belonging to
PT      Proteobacteria, gamma subdivision, etc.z
XX
PS      Claim 3; Page 5; 15p; Japanese.
XX
SS      The invention relates to a method for monitoring oil contamination of
CC      sea water by detecting, in the sea water, a microbe having the following
CC      properties: (1) it has no flagellum; (2) it is a Gram-negative bacterium;
CC      (3) it belongs to Proteobacteria, gamma subdivision; (4) it cannot
CC      assimilate glucose as a single carbon source; and (5) it efficiently
CC      assimilates at least one of 10-30C n-alkanes. This sequence represents
CC      a fragment of the 16S rDNA gene from the microbe of the invention,
CC      isolate SF-T1.
XX
SO      Sequence 1529 BP; 379 A; 351 C; 487 G; 312 T; 0 other;

Query Match      54.9%; Score 811.6; DB 20; Length 1529;
Best Local Similarity 76.2%; Pred. No. 3.1e-242;
Matches 1142; Conservative 0; Mismatches 314; Indels 42; Gaps 10;

Qy      1  gatcatgctcgaactaactcgtgcgcgtccttaacatgcaagtcgcggggtag 60
Db      2  gatcctgctcagattgaacgctggtgcgcgtcctaacaatgcaagtcgcggaaacg 61
Qy      61  caatc-----ctagcggcgaaagcggtgagtaacagctgtgaactcc 104
Db      62  atcctagcttgaagagcgctcgagcggtgcggtgagtaacagctgtgaactctcc 121
Qy      105  tcgagctgggataacttccgaagaagaactaataccgagatgctcgtgtgatacac 164
Db      122  catgagaggggataactccgtgggaacccaagcttaataccgataatccctaacgg 181
Qy      165  aagattgtaagtaagaattatgctgtggaatgagccgcgcgtatgactgaagtg 224
Db      182  agcaggggat--ctcgaccttgctgtagatgagtagcgtcgcgtcgtgattagctgt 239

```


PI	Kauplphen MS, Schuelein M, Schnotz K, Andersen LN, Bjornvad ME;	
DR	WPI: 2000-105891/09.	
XX	New mannanases for treatment of textiles, plant material and coffee	
PT	extract, and in cleaning compositions.	
XX	Disclosure: Page 242: 242pp; English.	
PS		
XX	The present sequence represents the 16S RNA sequence of <i>Bacillus</i>	
CC	sp. A112. This bacterium encodes a mannanase enzyme (also known as	
CC	mannan endo-1,4-beta-mannosidase, beta-mannanase, or endo-1,4-mannanase).	
CC	The mannanase hydrolyses galactomannans. Specifically, mannanases	
CC	hydrolyse 1,4-beta-D-mannosidic linkages in mannans, galactomannans,	
CC	glucamannans, and galactoglucomannans. The mannanase protein, or	
CC	preparations containing it, are used to improve properties of cellulosic	
CC	or synthetic fibres, yarn or (non)woven fabrics (removal of mannan-based	
CC	clays or printing pastes). They are also used to degrade or modify	
CC	plant materials (particularly recycled waste paper, paper making pulps,	
CC	or material containing guar or locust bean gums (thickeners), or to	
CC	reduce viscosity of mannan-containing foods or feeds). The mannanases	
CC	are also used to process coffee extracts (to inhibit gel formation); in	
CC	cleaning compositions (for machine washing of fabrics, as hard-surface	
CC	cleaners, for hard or machine dishwashing, also in oral, dental, contact	
CC	lens or body-care compositions) where they remove mannan-containing	
CC	soils and prevent binding of some soils to cellulose; and in fabric	
CC	softeners. They can also be used in oil well drilling to fracture	
CC	subterranean formations.	
CC		
XX		
SQ	Sequence 1508 BP; 384 A; 348 C; 464 G; 312 U; 0 other;	
	Query Match 54.9%; Score 810.2; DB 21; Length 1508;	
	Best Local Similarity 61.7%; Pred. No. 8,5e-242;	
	Matches 921; Conservative 220; Mismatches 303; Indels 49; Gaps 10;	
OY	15 actaacgcttgaggcgctcttaaacatgcaagctgagagc-----gtag 60	
DB	2 acgaacgcgagcgcgagcgcuaaaccgaugacgagcgagcgcauuaaggacuuugcucc 61	
OY	61 caatccctagcgcgcaacgcgagtgagtaaacctggtacatcttcctccg-agctgggata 119	
DB	62 uaaauuguaagcgcgcgagcgagugagguaaacacgucgggcaaccucuguaagacugggaa 121	
OY	120 acttcggaagaagactaacctcgcgagctcgtgttgatccacaagaattgatatgta 179	
DB	122 acaugagagaauacgugcuaauaccggaauaacuug-aggaugcuaauaccuuguaa 180	
OY	180 aagat-----tatgtctggagatgtagcccgcgccgcatgactgtgtgtgtagt 231	
DB	181 aagaaugcucggcuaauccuaacggaugagcgcgcgagcgcuaugacguuaguuagaa 240	
OY	232 aatgctccacaaagcgagacatcggtatagccgcgtgtagagggctgcgcacacatggaa 291	
DB	241 aacggcuaaccaaagcgacgagugcuaagcgcgcacugaggguguaucgycacacuggaa 300	
OY	292 ctgagacacgctcactactcctacgagagcgacagttaaagactcttgtaaatgggaa 351	
DB	301 cugagacacgcgcacagacucuaacgagcgagcgacgaguuagaaucuccgacaugaa 360	
OY	352 aaccttgaaagcagcgacgcgctgtagaagaagaagctcttgatgtgtaagtcatg 411	
DB	361 aagucugacgagcgcaacgcgcgucgaauguaauaaggguuucgucguuaaagcucug 420	
OY	412 gcaagaaaataagcagcatgtgtatgtgtgtaacctcgct-----aag 455	
DB	421 uuaagggaagaacaagayuccguucaaauagggcgcgccuugacguguaaccuaacagaag 480	
OY	456 caccgcttaactacgtgcgcagcagccgcggtatatacgtatgtgtcaagcgttgttgaa 515	
DB	481 ccaagcuaacuaacgucgacgagcgcguaauacguaugagugcagcguguuucggaa 540	
OY	516 tcatctggcgctaaagggtcgtagcgcgagtttgttaagtgatgtaaaactcgcggctc 575	

[illegible]

KX	Monitoring; oil; contamination; sea water; detection; flagellum;
KW	Gram-negative bacterium; Proteobacteria; glucose; carbon source;
KW	alkane; 16S rDNA gene; ds.
XX	
OS	Proteobacteria.
XX	
PN	JF11243967-A.
PD	
XX	14-SEP-1999.
PF	
XX	04-MAR-1998; 98JP-0069399.
PR	
XX	04-MAR-1998; 98JP-0069399.
PA	(KAIY-) KAIYO BIOTECHNOLOGY KENKYUSHO KK.
XX	
WI	WP1: 1999-564435/48.
DR	
XX	
PT	Monitoring of oil contamination of sea water - where oil contamination
PT	is evaluated by detection of a microbe having properties from e.g.
PT	having no flagellum, being a Gram-negative bacterium, belonging to
PT	Proteobacteria, gamma subdivision, etc.z
XX	
PS	Claim 3, Page 6-7; 15pp; Japanese.
XX	
CC	The invention relates to a method for monitoring oil contamination of
CC	sea water by detecting, in the sea water, a microbe having the following
CC	properties: (1) it has no flagellum; (2) it is a Gram-negative bacterium;
CC	(3) it belongs to Proteobacteria, gamma subdivision; (4) it cannot
CC	assimilate glucose as a single carbon source; and (5) it efficiently
CC	assimilates at least one of 10-30C n-alkanes. This sequence represents
CC	a fragment of the 16S rDNA gene from the microbe of the invention,
CC	Isolate Shm-2.
XX	
SQ	Sequence 1529 BP: 377 A; 352 C; 488 G; 312 T; 0 other:
Query Match	54.8%; Score 810; DB 20; Length 1529;
Best Local Similarity	76.2%; Pred. No. 9,8e-242;
Matches 1141; Conservative	0; Mismatches 315; Indels 42; Gaps 10.
OY	1 gatcatgctgcgaactaacgcgtcgccgccttcaacatgatcagtcgggtag 60
DB	2 gaccctggctcagatgtaacgcgtcgccgcctcaacacatgccaagtccgggaacg 61
OY	61 caatac-----ctagcggcgaaacgggtgtaatacacgtgtcaattccc 104
DB	62 atcctagcttgctagaggagcgctcgaacggcgagcggtgtgatacacgtgaatatctgcc 121
OY	105 tccgagttcgggtaactcttcgaaaaggaaagctaatacgcgatagtctgttgatcac 164
DB	122 catgtaggggggataacacctgggaaacccaggtcataacgcataatccctacgggggaa 181
OY	165 aagatttgataagtaaagatttatgcttggagatgagccccggccagattagctgttg 224
DB	182 agca--ggggacacctcggccttgctgtgtagatgtagctcgcgtctegattagctgttg 239
OY	225 gtbaagtaaatggtctcaaccaagcgcaagatgtgtagccgccttgagaaggggtgtccggccac 284
DB	240 gtbaagtaaatggtctcaccaaggcgacgatacttcgcttgagtgagatgatacagccac 299
OY	285 aatggaaactagacaagcgatccatactctcctacgggagagcaagcttaagaaattctgtcaa 344
DB	300 accgggaactagacaacggtcccggaactctctacgggaagcagtcagtgggaaattctgtgaa 359
OY	345 ttggggaaaccttgaaagcagcgacgcgcgttgaaacgaagaaggtctcggattgttaag- 403
DB	360 ttgggggaacacctgataccgacatgcgcgtgtgtgtgaagaagacgtcttcgggttgttaagc 419
OY	404 ----tcaattagcgacgaaaaataagcagcaatgtg-----atgatgttacctgct 451
DB	420 accttcagtttagggggaagagcttactcttataacagatgtagtactgtgacttaccataag 479

[illegible]

RESULT	13
AAx83565	
ID	AAx83565 standard; DNA; 1529 BP.

```
XX AA83565;
XX
XX 21-DEC-1999 (first entry)
XX
XX 16S rDNA gene fragment from marine bacterium isolate Wf-1.
XX
XX Monitoring; oil; contamination; sea water; detection; flagellum;
XX Gram-negative bacterium; Proteobacteria; glucose; carbon source;
XX alkane; 16S rDNA gene; ds.
XX
XX Proteobacteria.
XX
XX JPI1243967-A.
XX
XX 14-SEP-1999.
XX
XX 04-MAR-1998; 98JP-0069399.
XX
XX 04-MAR-1998; 98JP-0069399.
XX
XX (KAIY-) KAIYO BIOTECHNOLOGY KENKYUSHO KK.
XX
XX WPI; 1999-564435/48.
XX
XX Monitoring of oil contamination of sea water - where oil contamination
XX is evaluated by detection of a microbe having properties from e.g.
XX having no flagellum, being a Gram-negative bacterium, belonging to
XX Proteobacteria, gamma subdivision, etc.2
XX
XX Claim 3; Page 5; 15pp; Japanese.
XX
XX The invention relates to a method for monitoring oil contamination of
XX sea water by detecting, in the sea water, a microbe having the following
XX properties: (1) it has no flagellum; (2) it is a Gram-negative bacterium;
XX (3) it belongs to Proteobacteria, gamma subdivision; (4) it cannot
XX assimilate glucose as a single carbon source; and (5) it efficiently
XX assimilates at least one of 10-30C n-alkanes. This sequence represents
XX a fragment of the 16S rDNA gene from the microbe of the invention,
XX isolate Wf-1.
XX
XX Sequence 1529 BP; 376 A; 352 C; 489 G; 312 T; 0 other;
SQ
Query Match 54.7%; Score 808.4; DB 20; Length 1529;
Best Local Similarity 76.1%; Pred. No. 3,1e-241;
Matches 1140; Conservative 0; Mismatches 316; Indels 42; Gaps 10;
QY 1 gatcatgctcagaactaactcgtgctgctgtcttaacatgacgtcgcggtgag 60
DB 2 gatccgtgctcagatgacgctgctgctgctccttaacatgacgtcgcggtgag 61
QY 61 caatac-----ctagcgcgaacgggtgagtaacacgtgtgtaattcttc 104
DB 62 atcctagctgtcaggaagcgctgcagcgcggaaggtgagtaacgcgtgagaattctgc 121
QY 105 tccgaatcgtgataacttccgaagaaggaactaactacgcgaatgctctgtgtgacac 164
DB 122 cattatggtgggataaacctcggtggaacccaagcctaactacgcgaatgctctgtgtgacac 181
QY 165 aagatttgaatgaagaattatctgtctgtgagatgacccgcgcgcgaatgacgtgtg 224
DB 182 agca--ggggactctcgccctgtgtctgtatgtagtgacgtcgcgtctgtatgtgtg 229
QY 225 gtagagtaatgctcaccgaagcgacgacgtcgtgtacgcgcgcgcgaaggtgtgtcgcacac 284
DB 240 gtagagtaatgctcaccgaagcgacgacgtcgtgtacgtgtcttaagagtagtaacacacac 299
QY 285 aatggaactgagacacgtccatcaactcctacggaagcgacgaatgacgtctgtcaac 344
DB 300 accgggaactgagacacacgtccgaactcctacggaagcgacgaatgacgtctgtcaac 359
QY 345 tgggggaaacccctgaagcagcgacgcgcgtgaaacgaagaaggtcttcgattgtaaag- 403
```

```
DB 360 tgggggaaacccctgaagcagcgacgcgcgtgaaacgaagaaggtcttcgattgtaaagc 419
QY 404 -----ttcattagcaggaagaaataagacgaatgtg-----atgagttaacctgcct 451
DB 420 acttttagtaggaggaaggaagcttacttacttaacgaatgagtaactgtgactgtactacag 479
QY 452 a--aaagcaccgcttaactcgtgcagcagcgcggtlaactgattgtgtcgaacgctgt 509
DB 480 aagaagcaccgcttaactcgtgcagcagcgcggtlaactgattgtgtcgaacgctgt 539
QY 510 tgggaatcattggcgtlaaagggtgctagtcggaattgtlaagtcaggtgtgaaactgc 569
DB 540 tgggaatcattggcgtlaaagggtgctagtcggaattgtlaagtcaggtgtgaaactgc 599
QY 570 gggctcaaccgctgtgcctcacttgaactaactgaatctgtgagatttgggaaggaagctg 629
DB 600 gggctcaaccgctgtgcctcacttgaactaactgaatctgtgagatttgggaaggaagctg 659
QY 630 aattccaggtgtgagcgtgaaatgcgtagatctggaaggaacccaagtcgcaagcgca 689
DB 660 aattccaggtgtgagcgtgaaatgcgtagatctggaaggaacccaagtcgcaagcgca 719
QY 690 ctgtcgtgtcacaacacgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 749
DB 720 cctcctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 779
QY 750 ccccggttaactcagcgccttaacacgtgtctacacgtgtgtggtgtttaa--accctag 807
DB 780 cctcgtgtgtagcagcgccttaacacgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 839
QY 808 taacgaacttaacggttaagtagacgcctgtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 867
DB 840 tgaaggaacttaacggttaagtagacgcctgtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 899
QY 868 aaggaatctgacgggtgtccgcacaaagcggttgagacatgtgtgttaattcgtatgatacccc 927
DB 900 aatgaatctgacgggtgtccgcacaaagcggttgagacatgtgtgtgttaattcgtatgatacccc 959
QY 928 aaaaactcaaccgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 986
DB 960 aagaactcttaacggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1019
QY 987 --gcagattcaacaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1044
DB 1020 gggcgaagtgacacgggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1079
QY 1045 gtcccgcaacgagcgcaacccctatcgtatgtgc--taccttaagttggtgcaactgtgta 1101
DB 1080 gtcccgtaacgagcgcaacccctatcgtatgtgc--taccttaagttggtgcaactgtgta 1139
QY 1102 cgaactcgcgtgtgacaacacgggaagcggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1161
DB 1140 ggaactcgcgtgtgacaacacgggaagcggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1199
QY 1162 atgtccagtgcaacacgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1221
DB 1200 acggcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1259
QY 1222 gtagctaatctctaaagtcgtgtccagttcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1281
DB 1260 caagcaaatcccttaaaactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1319
QY 1282 agtcggaatcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1341
DB 1320 agtcggaatcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1379
QY 1342 tacacacgcgcgtcaccacacacacacacacacacacacacacacacacacacacacacacac 1401
DB 1380 tacacacgcgcgtcaccacacacacacacacacacacacacacacacacacacacacacacac 1438
QY 1402 aaggaagacgactaactgaagtcgtaaggggtgaaagtcgtaacgaagtaac 1459
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 01:16:45 : Search time 172.39 Seconds
(without alignments)
1940.414 Million cell updates/sec

Title: US-09-380-826A-1

Perfect score: 1477

Sequence: 1 gatcatgctcgcagactaac.....ccgtaatcgcttcctgcag 1477

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PT05.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	814.2	55.1	1536	US-08-642-229A-1	Sequence 1, Appli
2	810.2	54.9	1508	US-09-198-955A-14	Sequence 14, Appl
3	805.4	54.5	1521	US-08-501-126-20	Sequence 20, Appl
4	803.8	54.4	1536	US-08-114-695A-7	Sequence 7, Appli
5	801.8	54.3	1512	US-08-995-960-2	Sequence 2, Appli
6	801.6	54.3	1555	US-08-757-653-160	Sequence 160, App
7	796.4	53.9	1506	US-09-198-955A-13	Sequence 13, Appli
8	792.4	53.6	1517	US-09-248-528-2	Sequence 2, Appli
9	792.4	53.6	1517	US-09-549-108-2	Sequence 2, Appli
10	792.4	53.6	1517	US-09-549-111-2	Sequence 2, Appli
11	792.4	53.6	1517	US-09-549-106-2	Sequence 2, Appli
12	792.4	53.7	1517	US-09-550-394-2	Sequence 2, Appli
13	790.6	53.5	1500	US-09-193-377B-3	Sequence 3, Appli
14	789.8	53.5	1516	US-09-248-528-3	Sequence 3, Appli
15	789.8	53.5	1516	US-09-549-108-3	Sequence 3, Appli
16	789.8	53.5	1516	US-09-549-111-3	Sequence 3, Appli
17	789.8	53.5	1516	US-09-549-106-3	Sequence 3, Appli
18	789.8	53.5	1516	US-09-550-394-3	Sequence 3, Appli
19	781	52.9	1505	US-09-193-377B-5	Sequence 5, Appli
20	780.2	52.8	1542	US-08-114-695A-1	Sequence 1, Appli
21	780	52.8	1542	US-08-757-653-158	Sequence 158, App
22	779.8	52.8	1474	US-08-114-695A-8	Sequence 8, Appli
23	778.4	52.7	1556	US-08-995-960-1	Sequence 53, Appli
24	772.6	52.3	1484	US-08-632-470-53	Sequence 6, Appli
25	770.6	52.2	1518	US-08-114-695A-6	Sequence 2, Appli
26	769.6	52.1	1502	US-09-187-946-2	Sequence 27, Appli
27	768.4	52.0	1485	US-08-299-810A-27	

28	766.8	51.9	4403765	4	US-09-103-840A-2	Sequence 2, Appli
29	766.6	51.9	1542	2	US-08-875-445-21	Sequence 21, Appl
30	754.6	51.1	1494	2	US-08-632-470-49	Sequence 49, Appl
31	754.2	51.1	1452	2	US-08-642-229A-2	Sequence 2, Appli
32	753.6	51.0	1455	2	US-08-642-229A-3	Sequence 3, Appli
33	749.6	50.8	1464	2	US-08-938-858-1	Sequence 1, Appli
34	749	50.7	1508	2	US-08-632-470-44	Sequence 44, Appli
35	748.4	50.7	1428	4	US-09-193-377B-1	Sequence 1, Appli
36	747.8	50.6	1452	1	US-08-276-943-1	Sequence 1, Appli
37	747.8	50.6	1452	2	US-08-716-841-1	Sequence 1, Appli
38	745.8	50.5	1654	4	US-09-216-909-1	Sequence 1, Appli
39	745.8	50.5	1654	4	US-09-702-843-1	Sequence 1, Appli
40	745.8	50.5	1654	4	US-09-702-847-1	Sequence 1, Appli
41	744.6	50.4	1430	2	US-08-902-518A-1	Sequence 1, Appli
42	743.8	50.4	1415	2	US-08-632-470-52	Sequence 52, Appli
43	742.4	50.3	1503	4	US-08-943-571-1	Sequence 1, Appli
44	741.8	50.2	1440	2	US-08-632-470-26	Sequence 26, Appli
45	741.8	50.2	1441	4	US-09-193-377B-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1
US-08-642-229A-1
Sequence 1, Application US/08642229A
Patent No. 5874291
GENERAL INFORMATION:
APPLICANT: Herwig, Russell P.
APPLICANT: Bielefeldt, Angela R.
APPLICANT: Stensel, H. David
APPLICANT: Strand, Stuart E.
TITLE OF INVENTION: Degradation of Environmental Toxins by a
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Christensen O'Connor Johnson and Kindness PLLC
STREET: 1420 Fifth Avenue, Suite 2800
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: WA 98101-2347
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/642,229A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/246,865
FILING DATE: 20-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: Shelness, Diana K.
REGISTRATION NUMBER: 35,356
REFERENCE/DOCKET NUMBER: UOFW19233
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 682 8100
TELEFAX: (206) 224 0779
TELEX: 4938023
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1536 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: nucleic acid
DESCRIPTION: "16S ribosomal DNA"
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:

ORGANISM: Comamonas testosteroni ATCC No. 5874291 11996
US-08-642-229a-1

Query Match 55.1%; Score 814.2; DB 2; Length 1536;
Best Local Similarity 74.7%; Pred. No. 5.2e-284;
Matches 1113; Conservative 0; Mismatches 343; Indels 33; Gaps 6;

QY 1 gatacgtgctgaactaactacgtgagcgctgttaacatgcaagtcagcggtgag 60
DB 16 GATCCTGGCTCAGATTGAAAGCGTGGCATGCTTACATGCAATGCAACGTAA 75
QY 61 caa-----tacctacgagcgagcggtgagtaaccggtgtatctctccgag 110
DB 76 GGTCTTGGATGCTGACGAGTGGCGAAGCGGTGAGTATACATGGAACGTGCTAG 135
QY 111 tctggaataactcttcggaagaagaaactaatacagatgctctgttgatcacaagatt 170
DB 136 TGGGGATTAACACTACTCGAAGAGTAGTAATACCGCATGACATCTACGGATGAAGA 193
QY 171 tgaagtaagaattatctgttgagatagagcccgcgccgcatagctagtgtagg 230
DB 194 GGGGACCTTGGCGCTTGCTACTAGAGCGGCTGATGGCAGATTAGTAGTGTGGGG 253
QY 231 taatggtcacaagaagcgagcatcgtagccgctgagaggtgtccgagcacaatgga 290
DB 254 TAAAGGCTTACCAAGCGCTGATGTAAGTGTGAGAGAGACGACGACACTGGG 313
QY 291 actgagaacaggtcacaactctcagagagcgagcagttaaagattctgtccaatgggg 350
DB 314 ACTAGACACGCGCCAGACTCTCAGGAGGACACAGTGGGAATTTGGCAATGGGCG 373
QY 351 aaaccctgaagcagcgagcgctggaagaagaagtcctcgatgtgtaagttcata 410
DB 374 AAAGCTGATCAGCAATGCCGCGTGCAGATGAAGGCCCTCGGGTTGTAAACGCTTTT 433
QY 411 ggcagaaataaagca-----gcaatgtgatagtgtaactctgctcta--aag 455
DB 434 GTACGGAAAGCAAAAGCGGCGTAATATCCCGGTCATACCGTTACCAAGAAATAG 493
QY 456 caccggctaaactaactgctgcaagccggtgaatacgtatgtagcgaacgctgttcgga 515
DB 494 CACGGCTTAACCTAGCTGCGACAGCGCGGTAAATACGTAGGTCGAACGTTAATCGGAA 553
QY 516 tcaatggcgctaaaggggtcgtagcggaattgttaagtcagtcgtgcaactgcyggtc 575
DB 554 TTACTGGCGCTAAAGCGTGGCCAGCGGCTTTGTGAAGACAGTGTGAATCCCGCGCT 613
QY 576 aaccggtgctgcaactgtaaaactaacaagtcgtgagattggagagagcaagtggaattcc 635
DB 614 AACCTGGGAACCTGCCATTTGTGACTGCAAGGCTAGAGTGGCAGAGGGGATGAATTC 673
QY 636 aggtgtgaacgtgaaatgctgtagatatctggaagaacaacgtagcggaagcgacttgct 695
DB 674 GCGTGTACAGGAATCCGTAGATATCCGGAGAACCCGATGCGGAAGGGAATCCCT 733
QY 696 ggtcaaaactgaacgctgagcagcaaaagcgtggtagtaaaacggaattagatcccg 755
DB 734 GGGCTGACATGACCTCATGACAGAAAGCGTGGGAGCAAAACAGATTAGTACCTCGG 793
QY 756 taatcacgcctcaaaagctgtgtctaccagttgttgagggtttaaaccctcagtaacgac 815
DB 794 TAGTCCAGGCCCTTAAACGATGCAACTGTTGTTGTTAACTGACTCACTAACGAG 853
QY 816 ctaagcggttaagtagagcgcctgagggaactatgctcgcaagaagttaaactcaagaagt 875
DB 854 CTAACGCGTGAAGTGTGACCGCTGGGAGTACGGCCGCAAGTTAAACTCAAGCAATT 913
QY 876 gacgggggttcgcacaaagcggtgagcagctggtttaaactcagatgatacccaaaact 935
DB 914 GACGGGAGACCGCCACAAAGGGTGAATGATGTGTTAATTTCATGCAACGCAAAACCT 973
QY 936 caactggtgacatgagatctgaaatcagttagagatatatgagccttcgggcaga----- 991

DB 974 TACCACCTTTGACATGCGAGAACCTTACCAGAGATGTTGTGCTCGGAAGAAACCT 1033
QY 992 -ttcaagaagtgcagatggtgttcgcagactgctgcgtgagatgttgggttaagtcgg 1050
DB 1034 GCACACAGGTGTGCAATGAGGTGTCAGTCAAGTCTGTGTAAGTGTGGTTAAGTCCG 1093
QY 1051 caacgaagcgaaccccatcgatgtgttacccttaagtgtggcaactgagtaacagtcg 1110
DB 1094 CAACGAGCGCAACCTTGCCATTAGTTGCTACATTCAGTTGACACTCTAATGGAGCTGC 1153
QY 1111 cggltgacaaacgggaagagcggtgagatgacgtcaaatccatcagccttatgtccag 1170
DB 1154 CGGTGACAAACCGGAGGAAAGTGGGATGACGTCAATCCTCATGCGCTTAAAGTGGG 1213
QY 1171 gccacaagctgtcacaatgtccgatatcaagaggtgcgaaccccgcaagaagagctaat 1230
DB 1214 GCTACACAGCTCATACATATGGCTGTGTAAGAGGTTCCAAACCGCAGGGGAGCTAAT 1273
QY 1231 ctctaaagtcggtccagctgcatgtgggtgtctcaactcgaccccatgaagtgcgaat 1290
DB 1274 CCCATTAAGCCAGTCATGATCGGATGCGATGCAATCGACTGCGTGAAGTGGAT 1333
QY 1291 cgtatgaatcgcgagatcagatgacgagtcggtgtaactgctccggagactgtacaacgg 1350
DB 1334 CGCTAGTAATCGTGTGACAAATGTCACGGTGAATACGTTCCGGGTTGTACACACG 1393
QY 1351 cccgtcacacacctgagtggtggagcgacccggaagtgtgtctgttgaacgtaagagaca 1410
DB 1394 CCCGTCACCATGAGGAGCGGCTCTCCGCAAGATAG-GRAGCCTAACCGTAAGAGAGGC 1452
QY 1411 gactactaagtgtaaacctcgtaaaagggtgtaagtcgtaacaagttacc 1459
DB 1453 GCTTACCAACGCGCGGTTGCTGACTGGGTCGTAACGTCTACCAAGTAGC 1501

RESULT 2
US-09-198-955a-14
Sequence 14, Application US/09198955A
Patent No. 6187580
GENERAL INFORMATION:
APPLICANT: Andersen, Lene N.
APPLICANT: Schlein, Martin
APPLICANT: Lange, Niels E.
APPLICANT: Bjornvad, Mads E.
APPLICANT: Moller, Soren
APPLICANT: Glad, Sanne O. S.
APPLICANT: Kaupinen, Markus S.
APPLICANT: Schmitt, Kirk
TITLE OF INVENTION: No. 6187580el Pectate Lyases
FILE REFERENCE: 5378.200-US
CURRENT APPLICATION NUMBER: US/09/198, 955A
CURRENT FILING DATE: 1998-11-24
PRIOR APPLICATION NUMBER: 1343/97
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 1344/97
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/067, 249
PRIOR FILING DATE: 1997-12-02
PRIOR APPLICATION NUMBER: 60/067, 240
PRIOR FILING DATE: 1997-12-02
PRIOR APPLICATION NUMBER: 09/073, 684
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 09/184, 217
PRIOR FILING DATE: 1998-11-02
NUMBER OF SEQ ID NOS: 32
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14
LENGTH: 1508
TYPE: RNA
ORGANISM: Bacillus sp.
US-09-198-955a-14

Query Match 54.9% Score 810.2; DB 4; Length 1508;
 Best Local Similarity 61.7%; Pred. No. 1.4e-282;
 Matches 921; Conservative 220; Mismatches 303; Indels 49; Gaps 10;

```

15 actaacgtgctgctgctgctttaaactgcaagtcgagcgg-----gtag 60
   |||||||:|||||:|:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2 acgaagcugcgcgcgugcuaaauacugcaagcgagcgagacuaauagagcuuucc 61
   |||||||:|||||:|:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 caatactgacgagcgaagcggtgtagtaacaagtgtatcttcccg-agtctggata 119
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 62 uaaauuagcgcgcgagcgugagguuacacugyggcaacugccuugagacuggana 121
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 120 acttccggaaggaagcgaatactacgtagtctgttgatcacaaagtttgatgta 179
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 122 acucagaaauugugucuaaauaccgaaauauuug-aggauugcuaauccuuuua 180
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 180 aagat-----tattgttggagatgagccgagcgagcttaagtgtgtagt 231
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 aagaugugucugcguaucacugagguagggccgagcgagcuaugcuuguguaagu 240
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 232 aatgtctacacaaagcgagcagatcggtagccgagctgagaagtggtccgcacaaatgaa 291
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 241 aacguguaacaaagcgagcgaugcguaugcgacgagagggugauucgagcacauggaa 300
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 292 ctgagacaggtctacatctacgagagcgagcagtaagaatctgtctcaatgggga 351
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 301 cugagacacgccccagacgctcuaagcgagcgagcagaaugaggaauuucgcaauagacga 360
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 352 aacccgaaagcagcgagcagcgctgagcagaaaggtctctgagatgttaagcttag 411
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 361 aagucagagcagcgagcagcgcgugagugagagggguuucgucuguaagcucugung 420
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 412 gcaggaataaataagcagcaatgtgtgtgtctgct-----aaag 455
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 421 uaaagggaataacagaugcgguucaaauagggcgacacuuagcguaacuaacagaag 480
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 456 caccgctactactgctgacgagcagcgagtaactgtagtgcgaagctgttgagaa 515
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 481 ccaagcguaacuaugcugcagcgagcgagaaacuaacuaugugcgagcuuucgga 540
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 516 tcaatggcgtaaaaggtgcgtgagcgagcttgtaagtcaggtgtaaaactgcgggctc 575
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 541 uuaugugcguaaagcgcgagcgagcguguuuuaugucugagugaaucucggggcuc 600
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 576 aaccgctgctgctgacttcaactacagtcgtgagcttgggagagcgaaagtgaattcc 635
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 601 aaccccgagcguguaacuaugagaaacugagaaacuaagagagagagugaaaucc 660
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 636 aggtgtgagcggtgaaatgctgtagatactctgaggaacacagagtcggaagcgagctgtc 695
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 661 acgugagcguguaaagcguaaauugugagaaacacagugugcgaaagcgagcucucu 720
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 696 ggtctcaaacctgacgctgagcagaaagcggtggtgtagtaaacggtgataaccgg 755
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 721 gguucuguaacugcgugagcgcgagaaagcgugagcgaaacagaaauuagauccug 780
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 756 taatccagcgcccaaacgctgtctaccagtgcttggtgtttaa--ccctcgtatcga 813
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 781 uaguccaagcguaaagcguaagugcuaagguuagggguuucgagcccuuagucgca 840
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 814 actaacggaataagtagagcgcctgaggaactgtctgcgaagagtgaaactcaagaa 873
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 841 aguaaacaacuaaagcacucgcugggaggaagcagcagcaagguuagaaacucuaagaa 900
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 874 ttagcgggggtcccgcaaacggttgagacatgtggtttaaictgatgatacccaaaaac 933
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 901 uugacggggggcccgcaaacgagugagcagugguuuaauucgaaacgagcaagaaac 960
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 934 cttaactggtgctgatactgatactgatactgtagagata--tatgagccttgaggaga 991
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 961 cuuacagugucuaagacucuaagacucuaagagagagagauuucuuucgggggaca 1020
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

QY 992 t-----tcacaagtgctgcatgtgtctgcacagctgctgtcgtgagatctgggttaagtc 1048
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1021 uaauguaacagugugcuaugugucugacugcugcugugagagauugugguuaagucc 1080
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1049 cgcacaagcagcgaacccct-atcgtatgtgtctacttaagtgtggcactggtacgaac 1107
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1081 cgcacaagcagcgaaccccuuagucuaugugcagcagcauuuugggcgacucuaagugac 1140
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1108 tgcggtgacaacacggaggaagcggtgagatgacgtcacaatctcatgtgcttatgtcc 1167
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1141 ugcguguguaaaccgaggaagugggagagcagcuaacucacugcuccuuaagacac 1200
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1168 agggccacacagctgtctactatgctacagaggggtgcacacgcgaagaaggagac 1227
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1201 uggucacacacagucacaaauagauuguaacaaagcagcaaacgcgagagugagcc 1260
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1228 aatctctaaagtcggtlcccaagtcgagatgtggtgtctgcaactgcagcccatgaagtcgg 1287
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1261 aauucuaaagcacaauuucuaugucgaguuugagugagcugcacaucgcauagaaagcg 1320
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1288 aatcgtatgataatcgcgagatcagatcgcggtgagatgataatgtccggactgtatcaca 1347
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1321 aaugucuaagaaucgagagcagcagcagcagcagcagcagcagcagcagcagcagc 1380
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1348 ccgcccgtcacacacactgtggtgagagcagccgaagtgtctgtttaaactga-agg 1406
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1381 ccgcccgtcacacacactgtggtgagagcagccgaagtgtctgtttaaactga-agg 1440
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1407 gacagactactaagtgtaaacctgctaaaggggtgaaagtgtgtaacaaagtacc 1459
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1441 gcuagccgcuaaagugggagcagagauugugggugagugcuaacaaagtacc 1493
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

RESULT 3
US-08-501-126-20
; Sequence 20, Application US/08501126
; Patent No. 6140095
; GENERAL INFORMATION:
; APPLICANT: Van Solingen, Pieter
; APPLICANT: Williams, Diane P.
; APPLICANT: Iverson, Sara
; APPLICANT: Farrell, Roberta L.
; APPLICANT: Herbes, Wilhelmina T.
; APPLICANT: Van Der Kleij, Wilhelmus A.
; APPLICANT: Herweijer, Margaretha A.
; APPLICANT: Van Beckhoven W.C., Rudolf F.
; APPLICANT: Jones, Brian E.
; TITLE OF INVENTION: ALKALI-TOLERANT XYLANASES
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 2000 Pennsylvania Avenue, NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/501,126
; FILING DATE: 29-DEC-1995
; CLASSIFICATION: 425
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 4615-0057.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500

```

TELEFAX: (202) 887-0763
: INFORMATION FOR SEQ ID NO: 20:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1521 base pairs
: type: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: HYPOTHEetical: NO
: ORIGINAL SOURCE:
: ORGANISM: Bacillus sp.
: INDIVIDUAL ISOLATE: DSM 8721
: US-08-501-126-20

Query Match 54.5%; Score 805.4; DB 3; Length 1521;
Best Local Similarity 76.9%; Pred. No. 7.9e-281;
Matches 1097; Conservative 0; Mismatches 292; Indels 37; Gaps 8;

QY 69 agcgagcgaacggtgagtaacacgt-ggtaattctctccgaagctggataactccg 127
DB 84 AGCGCGGACGGGTGAGTACACGCGGCAACCTTGTAGACTGGATACCTCCGGG 143
QY 128 aaagaaagcgaacacgagatcctgctgagacacaaattgatagtaaatla 187
DB 144 AAACCGGGGCTAATACCGGATGATCAT-TTGGATCGCATGATCCGAATGTAAGTGGG 202
QY 188 tgg-----cttgagatgagcccgcgcgagtaagtaagtgtgaggtatagc 237
DB 203 ATTATTCCTCACATCGACGATGAGTGGCCGCGCCGCTTACCTAGTGTGATGATGATG 262
QY 238 tcaacaagcgaacgagatcgtgagcgcctgagaggtgtccgcccacaatgaaactg 297
DB 263 TTACCAAGCGCGACGATGCTGACCCACCTGAGAGGGGTATCGCCACACTGGAACTG 322
QY 298 cagcgttccatctcctcgaagcgaacgaatgaatctgtctcaatggtgggaaacct 357
DB 323 CACGGCTCAGACTCTCTACCGGAGGAGCAGTGGGATCATCCGGAATGGGGAAGCT 382
QY 358 gaagcagcgaacgacgctgtaacgaagaaggtctcgagatgttaagttataggcga 417
DB 383 GACGGTCAACGCGCGCTGACGATGAGATGAGTTTTCGATCGTAAGTTCTGTTATGAGG 442
QY 418 aaataagcgaacatgtgatagtgatactgct-----aaagcagcg 461
DB 443 AAGAACAAGTGGCGTTCGATAGTGGCAGCTTGACGCTACCTCAGGAAGACCCCG 502
QY 462 ctaactcgtgacgaacgacgctgtaatactgataagcgtgtgtcgaatcatg 521
DB 503 CTAACTACGTCGACGACGCGCTCTTAAGTGTGATGTAAGGCCACAGGCTCAACCG 562
QY 522 ggcgtaaagggtgctgagcgtgattgtaagcaggtgtaaaactcggtcaacccg 581
DB 563 GCGCTAAAGCGCGCAGAGCGGCTCTTAAGTGTGATGTAAGGCCACAGGCTCAACCG 622
QY 582 tggcctgacttgaaactacaagtcgtgagtttggagaggaagtgaaatccagtg 641
DB 623 GAGGCGTCAATTTGAAACTGGGAGCTTGAGTGTGAGAGGAAATTCACAGTGT 682
QY 642 agcggtgaaatgtgatactctgaggaagacacacagtgaggaaggaactgtgctca 701
DB 683 AGCGGTAAATGCTGATATGTTGAGAGAACACCAAGTGGGCAAGGCACTTTTGCCCTA 742
QY 702 aaactcgtgagcgaacgaacgtgtgagtaaaacggaattagatacccggtatcc 761
DB 743 CAACGTGACGTGAGCGCGCAAGCGTGGAGCAACAGATTAAGATACCTGTTAGTCC 802
QY 762 acgcctaaacgtgtctacacagttgttgggtt--taacctcagtaacgaactaa 819
DB 803 ACGCCCTAAACGATGAGTGTGATGTTAGGCTTTGATGATCCCTTAAGCCGAGATTA 862
QY 820 cggagtaagtagaccgctcgaggaactatgctcgaagaagtgaactgaagatgacg 879

DB 863 CACATTAAAGCACTCCGCTGGGAGTACGCGCAAGGCTGAACCTCAAGAATTTAGC 922
QY 880 ggggtccgcaaacggtgagacatgtggtttaattcgaatgaatacccaaacctcacc 939
DB 923 GGGCGCCGCAAGACAGTgagacatgtggtttaattcgaatgaatacccaaaccttacc 982
QY 940 tggcctgaca-----tggatcgaatcatgtagagatatagaccttcggcagatcc 994
DB 983 AGGCTTTGACATCTCTGACACCTTGAGACAGACGCTTCCCTTGGGGACACAGTGT 1042
QY 995 acaggtgctgacatgtgtgtcgtgaagctgtgtgtgagatgttgggttaagtcggcaac 1054
DB 1043 ACAGGTGTGATGTTGCTGATGCTGCTGTGATGATGTTGGTTAAGTCCCGAAC 1102
QY 1055 gaggcgaacccct-atcgatgtgtctgaaccttaagttgagcactggtacgaactgccc 1113
DB 1103 GAGCGCAACCTTGATCTTGTGTTGCCAGCTTTCAGTTGGGCACTTAAGTGTACCTCCG 1162
QY 1114 tgaacaacggaaggaacggtgagatgacgtcaaatcccaatgaccttattgtccagggc 1173
DB 1163 TGATTAACCGGAGGAAGTGGGATGACGTCAATCATCTATGCCCTTATGACCTGGGCT 1222
QY 1174 acacagctgctacatgtgacgatacagaaggtgtccgaactcgaagaggaagtaactc 1233
DB 1223 ACACAGCTGCTACATGATGATGTACAAAGGCAACGACGACCGGAGGTTAAGCATCC 1282
QY 1234 taagaagtcgtccagctcgtgattggtgtctgcaactcgaacccatgaagtgcgaatcgc 1293
DB 1283 ATTAAGGCAATCTCAGTTGGATGTCAGGCTGCAACCTGCTGATGATGACCGGAATTC 1342
QY 1294 taagtaactcgtgacatgacatgacggtgagtaagttccggaactcgtgacacacgccc 1353
DB 1343 TAGTAATGCGCGATGACATGCGCGGATGATGATGTTCCGAGGCTTGTACACACGCC 1402
QY 1354 gtcaacacccctgagtggtgagcaccggaagtgtcttgttlaacglaaggaagaac 1413
DB 1403 GTCAACACGAGAGTGTGTAACACCGAAGTGC-GTGGCGTAACTTTGGACCGACC 1461
QY 1414 tactaagtgaaactcgttaagaggtgtgaagtgtgtaacgaaggtacc 1459
DB 1462 GNGCAAGTGGGACAGATGATGGGTGAAGTGTGTAACAGATATC 1507

RESULT 4

US-08-114-695A-7

: Sequence 7, Application US/08114695A

: Patent No. 5508193

: GENERAL INFORMATION:

: APPLICANT: Mandelbaum, Raphael T.

: APPLICANT: Mackett, Lawrence P.

: TITLE OF INVENTION: DEGRADATION OF S-TRIAZINES IN SOIL AND

: TITLE OF INVENTION: WATER

: NUMBER OF SEQUENCES: 8

: CORRESPONDENCE ADDRESS: LUNDBERG & WOESSNER, P.A.

: ADDRESS: 3500 IDS CENTER

: CITY: MINNEAPOLIS

: STATE: MN

: COUNTRY: USA

: ZIP: 55402

: COMPUTER READABLE FORM:

: MEDIUM TYPE: Floppy disk

: COMPUTER: IBM PC compatible

: OPERATING SYSTEM: PC-DOS/MS-DOS

: SOFTWARE: Patent Release #1.0, Version #1.25

: CURRENT APPLICATION DATA:

: APPLICATION NUMBER: US/08/114,695A

: FILING DATE: 31-AUG-1993

: CLASSIFICATION: 435

: ATTORNEY/AGENT INFORMATION:

: NAME: MUEITING, ANN M.

: REGISTRATION NUMBER: 33,977

: REFERENCE/DOCKET NUMBER: 600.268051

```

TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-339-0331
TELEFAX: 612-339-3061
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1536 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: RNA
ORIGINAL SOURCE:
ORGANISM: Pseudomonas testosterone
US-08-114-695A-7

Query Match      54.4%; Score 803.8; DB 1; Length 1536;
Best Local Similarity 60.2%; Pred. No. 3e-260;
Matches 897; Conservative 216; Mismatches 343; Indels 34; Gaps 7;

OY 1 gatactgctgaactaacgcttgccgagc-gcgtcttaactgcaagtcgagcg-55
DB 15 GATCCGCGCAGAUUGAAGCGCGGCAUNGCUAACAGCAAGCGAAGCGUAAC 74
OY 56 -----gtagtaactactagcgcgcgaacggttgagtaacagctgttaactctccga 109
DB 75 AGGCUUCGAGUUGAGCAGGCGAAGCGGUGAGUAUACUACGAGCGUAGCUAGUA 134
OY 110 gctcgtgataactctcgaagaagtaactacggatagctcgttctgtatcacaagt 169
DB 135 GGGGGGAAUACUACUAGAGUAGCUAAUACCGCAUAGAUUCUACGAGAAAGCA- 193
OY 170 ttagatagaaagatttatctgttgagatgagcccgagccgattagatgttgtag 229
DB 194 -GGGACCUUCGCGCCUUCGUCACUAGAGCGCGUAGUGGCAUAGUAGUGUGG 252
OY 230 gtaactgctcaacgaagcgatcgttagccgctgagaggtgtccgacacatg 289
DB 253 GUAAAGCGUACCAAGCCUGCAGUUCUGUAGUGUAGAGAGACAGCACAGCAGCUG 312
OY 290 aactgagaaacgctcactactcctacggagagcagcttaagaatctgtccaatgag 349
DB 313 GACUAGACACAGCCGACAGCUCUACGGGAGCAGCAGGCGGAUUAUUGGACAAUGGCG 372
OY 350 gaaccctgaagcagcgcgctgaaagaaaggtcttcggatgttaagttcatt 409
DB 373 GAAAGCCUUAUCCAGCAUAGCCGCGUGCAGAUAGAGCCCGCGGUGUUAACUCUU 432
OY 410 aggcagaaataagcag-----caatgtatgtatgttacctgcta--aa 454
DB 433 UGUACGGAACGAAAGCCUGGCGCUAAUUAUCCCGGUCUAGACGUAACGUAAGAUA 492
OY 455 gaccggctaaactgctgcaagcgcggttaactatgctgtgaacgctgttcgga 514
DB 493 GCACCGGCUAACUACGUGCCAGCAGCCGCGUAUUCGAGGUGGCAACCGUUAUACGGA 552
OY 515 atcattgagcgttaaggtgctgtagcggatgtgaatcagtggtgaactgcgagct 574
DB 553 AUUACUGGCGGUAAGCGGCCAGCGGCUUUUGUAAGACUAGUGGAUUAUCCCGCGCU 612
OY 575 caaccgctgagccttgaaactgaactgaagcttgaggttgagagagcgaagtgaatc 634
DB 613 CAACCGGGAACUGCCAUUGUAGUAGCAGGCUAGAGUGCGCAGAGGGGAGUAUUC 672
OY 635 caggtgtgagcgttgaatgctgtagatactggaagaaacccagtgcggaagcgactgc 694
DB 673 CCGGUGUACACUAGAUUUGAGAUUCCGAGAAACCGAUGCGCAAGCAUACCC 732
OY 695 tggctcaaaacgacgctgagcagcaagcgctggttagtaacggatagataaccgcg 754
DB 733 UGGCGCUGCAGCAGCCUUAUCCAGCAAGCGUGGAGCAAAACGAGAUUAGUACCCUG 792
OY 755 gtaatcacgacctaaagctgtctacagctgttggtggttttaaccctcagtaacga 814

```

```

DB 793 GUAGUCCACGCCCCUAAACGAGUUCACUGGUUGUGCUUAACUGACUCAGUAACGA 852
OY 815 cctaacgattaaagtagaccgctggggaactgctcgcaagtgaaactcaagaagt 874
DB 853 GUAAACGCGUAGAUUGAGCGCGCGGAGUACGCGCCGCAAGCGUUAACUCAAAGGAU 912
OY 875 tgaaggggttcgcacaacggttgagcagtgtgttaattgaatgatacccaaac 934
DB 913 UGAGGGGACCCGCAACAGCGGUGAGUAGUUAUUCAGUACAGCAAGGAAC 972
OY 935 tcaactggcttgacatgagatcgtatcaatgagataatgagcctcgggcaga--- 991
DB 973 UUAACCCACCUUAGUGCAGAGAAUUAUAGUUGUUGUGUCUGGAAGAAAC 1032
OY 992 ---ttacaggttcgactgctgctgacgctgctgtgtgagatgtgtgtaagctcc 1049
DB 1033 UGCACACAGUUCUGAUUGGUGUCCGACGUCGUGUCUGAGAUUUGGUAUACUCC 1092
OY 1050 gaaacgagcgaacccctatgtatgttgaactgaattggaactgtagcaaatg 1109
DB 1093 GCACGAGCGCAACCCUUCGCAUUAUAGUUCUACAUUACAUAGCACUCUUAUAGGACUG 1152
OY 1110 ccggtgacaaacgcgaggaagcggtgagatgacgtcaatcctcatgttctcag 1169
DB 1153 CCGGUGACAAACCGAGAGGAGGUGGAGUAGACUGCAUUCUACUAGGCCCUUAUAGUGG 1212
OY 1170 ggcacacacgtgtcaatgtgcgatacagaaggttcgccaactgcgaagggacta 1229
DB 1213 GCGUACACAGCUCUACAUUAGUGUGUACAAAGGUGUGCAACCCCGAGGGGACUUA 1272
OY 1230 tctctaaagctggtcccgatgagatgttggtctgcaactggaaccctgaagtcgga 1289
DB 1273 UCCCAUAAACCGAAGUGUGUCCGAGUCCAGUCUACUCCAGCUGCCUUAAGUGGA 1332
OY 1290 tgcgtgaatcgcgagatcagatcgcggtgagatgctccggaactgttacacac 1349
DB 1333 UCGCUUAUAGUGGAGUACGAUUGUACGUGCAUUAUAGCUGCCGCGUUCUAGACACC 1392
OY 1350 gccgtcaacacacactgagttgggagaccccgaaagtgtgtcttgaaccgtgaagac 1409
DB 1393 GCCCGCACACCAUUGGAGCGGUGUCGCCAGAGAAG- GUAGCCUAAACCGUAAGGAGG 1451
OY 1410 agacttaagatgaaactcgtaaaggggtgaagtcgtgaacgaagttaac 1459
DB 1452 CCGUUAACAGCGCGGUGUUCUACUGGUGUAGUAGUUAACAGAGUAC 1501

RESULT 5
US-08-995-960-2
Sequence 2, Application us/08995960
Patent No. 6030818
GENERAL INFORMATION:
APPLICANT: PAGI, Michel
APPLICANT: LANDRY, Nathalie
APPLICANT: BOISSINOT, Maurice
APPLICANT: HILIE, Marie-Claude
APPLICANT: HARVEY, Mario
APPLICANT: GAGNI, Martin
TITLE OF INVENTION: BACTERIAL MASS PRODUCTION
TITLE OF INVENTION: OF TAXANES AND PACLITAXEL
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSER: Dike, Bronstein, Roberts & Cushman, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: U.S.A.
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FastSeq for Windows Version 2.0b

```

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/995,960
FILING DATE: 22-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: RESNICK, David S
REGISTRATION NUMBER: 34,235
REFERENCE/DOCKET NUMBER: 1826/47986
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1512 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: rRNA
US-08-995-960-2
```

Query Match 54.3%; Score 801.8; DB 3; Length 1512;

Best Local Similarity 75.3%; Pred. No. 1.6e-279;

Matches 1134; Conservative 0; Mismatches 322; Indels 49; Gaps 9;

```

OY 1 gataatgctagaactaagctggtggtggtttaaacaatgcaagtcgagggtag 60
    |||||
DB 8 GATCATGCTCAGATGAAACCTGGCGGCTGCTTAATGCAAGTCGAGCAATGGA 67
OY 61 caat-----acctaagcgagcgaggtgtagtaacagt-gtaatcttc 103
    |||||
DB 68 TTAAGAGCTTCTCTTATGAAGTTAGCGCGGAGCGGTAGTAACCTGGTAACTAC 127
OY 104 ctccgagctctggaataactcttcgaaagaaagtaataccgtagtctgttgaatca 163
    |||||
DB 128 CCATTAAGACTGGATTAACCTCGGGAAACCGGGGCTAATACCGGATTAATTTGAACTGC 187
OY 164 caagattgataagtaag-----attatgcttgagatgagcccggtcgcatata 216
    |||||
DB 188 ATAGTTCGAATTTGAAGACGGCGCTTGCTGCTTATGATGATGACCGCGCTGCATTTA 247
OY 217 gctagttgtagtgaatgagctcaaccaagcgcaagatcgtagccggtctagaaggtgt 276
    |||||
DB 248 GCTAGTTGGTGGATGAGGCTCACCAAGGCGAGCATGCGTAGACCGGACTTAGAGAGGTGA 307
OY 277 ccgagcacaatggaacttgagacaacggtccatactctcaacggaagcgagcagttaaagaatc 336
    |||||
DB 308 TCGGCGACACTGGGACTGAGACACGGCCAGACTCTCAACGGGAGGACAGTAGAGGATTC 367
OY 337 ttgtcgaatggggaacccctggaacgagcgagcggtgtagaagaaaggtctcgat 396
    |||||
DB 368 TTCCCAATGAGCAAGCTGTACGAGCAACCGCGGTGATGATGAGGCTTCGGGT 427
OY 397 tgaagtctcatagcaggaataaagcaagcaatgtgatgatgtacgtcct----- 451
    |||||
DB 428 CGTAAACTCTCTTTTGGGGAAGAAAGATGCTAGTTGAATAAAGCTGGACACTTGACGG 487
OY 452 -----aaagcaacggttaactacgtagcagcaag-ccggtgtaatacgttgtgtg 499
    |||||
DB 488 TACTTAACCAAGAACCAACGCTTAACGTACGTCCAGACCGCGGTATATACGTAGGTGG 547
OY 500 caaggtgtctggaatcatcttggtggttaaggtgtgtagaggtgttgaagtagtga 559
    |||||
DB 548 CAAGGCTTATCCGGAATTAATGGCGTAAAGCGCGGCGAGGTGTTCTTAAGTCTGATG 607
OY 560 tgaaacctgcggaacccgctgagccttgaaactgaactacagatctggaatcttggag 619
    |||||
DB 608 TGAAGCCCAACGGCTCAACCGTGGAGGCTCATTTGGAAGAACTGGGAGAGCTTGAGTGAAG 667
```

```

OY 620 aggcaagtgaattccagtgtagcggtgaatctgtagatctctgtaggaacaccagt 679
    |||||
DB 668 AGGAAGTGGAAATTCATGATGCTAGCGGTGAATGCGTAGAGATATGAGAGAACCAAGTG 727
OY 680 gcgaaaggcagcttctgctcaaaactgacgtcgaggaacgaagcggtgtgttaag 739
    |||||
DB 728 GCCAAGCGCACTTCTGCTGTAACTGACATGAGCGCGGAAGCGTGGGGAACAA 787
OY 740 ggaatgatacccccgtgaatccagccctaaacgttcttaccagttgtggtgttt- 798
    |||||
DB 788 GGATTAGATACCTCTGATGTCACGCCGTAAAGATGATGATGATGATGAGGGTTTC 847
OY 799 -aacctcagtaacgaacctaagatagtagaccgctggggactatgctcgcaaga 857
    |||||
DB 848 CGCCCTTATGCTGAAGTTAAACGATTAAGCACTCCGCTGGGAGTACGGCCGAAG 907
OY 858 gtgaactcaagaaatgagcggtgcgcacaaagcggtgagcatgtgtttaaattcg 917
    |||||
DB 908 CTGAAGCTCAAAAGAAATGACGGGGGCCGACAAAGCGGTGAGCATGTGTTAAATTCG 967
OY 918 atgatcccccaaaacctcaactgagcttgacatgtagtctgaaatcaltagagatatag 977
    |||||
DB 968 AAGCAACGCCAAGAACCTTACACAGCTCTTGACAT-CCTGTGAAGAACTTAGAGATAGC 1026
OY 978 agc-----ctcgggcaagatcaacagtgctgcatggtgtgtcagctgtgtgtgaga 1033
    |||||
DB 1027 TTCTCTCTCGGAGCAAGTGAAGGAGTGCATGATGTTGCTGATGCTGCTGCTGAGA 1086
OY 1034 tgttggttaagtcgcgcaacgagcgcaacccct-accgtatgttcttaacttaagtgtg 1092
    |||||
DB 1087 TGTGGGTGAAGTCCCGCAACGACGCAACCTTGATCTTAAGTGGCATCATTAATTAAGTTGG 1146
OY 1093 gcactggtlaacgaactgcgcgttgacaacacggaaggaagcggtgtagtcaatcttc 1152
    |||||
DB 1147 GCACCTTAAGGTGACTGCGGTGACAAACCGAGAGAGTGGGATGATGATGATGATGATC 1206
OY 1153 atgacctttagtccagagcgcaacagctgtacaaatgagcggtgagagaggtgcgaac 1212
    |||||
DB 1207 ATGCCCTTATGACTGGGCTACACAGTGTGCACAAATGAGAGGTACAAAGCTGCAAGA 1266
OY 1213 tcgcaagaggaagctaatctctaaagtcggtccagttcggatctggtgtgtaactcg 1272
    |||||
DB 1267 CCGCGAGGTGAGGCTAATCTATTAACCCGTTTCAGTTCGATTTGAAGTTCGACATCGC 1326
OY 1273 acccatgaagtcggaatcgtagtaacggtacagatcgagtcggtgtgtaactcttc 1332
    |||||
DB 1327 CCTACATGAAGCTGGAATCGCTAGTAATCGCGATACAGCATGCGCGGTGAATACGTTCC 1386
OY 1333 cggacctgtgtaacacgccccgtcaacaccctgaagtggtggaagcaccggaagtgtctt 1392
    |||||
DB 1387 CGGGCTTGTACACACCGCCGCTCACACAGAGATTGTAACACCCGAAGTGGGTGG 1446
OY 1393 gtaacgctlaaggaagacagactactaagtgtaaaactcgtaaaaggggtgtaagtcgtaca 1452
    |||||
DB 1447 GTAACTTTATGAGACCCAGCGGCTTAAGGTGGAGCATGATGTGGGTGAGTGTGATACA 1506
OY 1453 aggtta 1457
    |||||
DB 1507 AGGTA 1511
```

```

RESULT 6
US-08-757-653-160
; Sequence 160, Application US/08/757653
; Patent No. 5843669
; GENERAL INFORMATION:
; APPLICANT: Kaiser, Michael W.
; APPLICANT: Lyamichay, Victor I.
; APPLICANT: Lyamichay, Natasha
; TITLE OF INVENTION: Cleavage Of Nucleic Acid Using
; TITLE OF INVENTION: Thermostable FEN-1 Endonucleases
; NUMBER OF SEQUENCES: 190
; CORRESPONDENCE ADDRESS:
```

```
ADDRESSER: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/757,653
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02565
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEO ID NO: 160:
SEQUENCE CHARACTERISTICS:
LENGTH: 1555 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-757-653-160
```

```
Query Match      54.3% Score 801.6: DB 2: Length 1555;
Best Local Similarity 75.3% Pred No. 1.9e-279;
Matches 1134: Conservative 0: Mismatches 324: Indels 48: Gaps 9;
```

```
QY 1 gatcatgctcgaacgaacacgctgagcgagcgtcttaaacatgcaagcgaggtag 60
DB 16 GATCGTGGCTGAGATGAGACGGTGGCGGCTGAATACATGCAAGTCGAGCGA 75
QY 61 caatacc-----tagcgcgaaacggtgtagtaaacgctgg-taaccttc 103
DB 76 CGAAGGCTTGTCTCTGATGTTAGCGGGGAGCGGTGAGTAACCGTGATTAACCTAC 135
QY 104 cccgagcttgatgaactcttcgaaagaaagataaccgagatgctcgttgatca 163
DB 136 CATATAGACTGGGATTAACCTTCGGGAACCGGACGTAATACCGGATAT-ATTITTAACCG 194
QY 164 caagattgatagttaaaga-----ttatctgtgagatgagcccgcgagcga 216
DB 195 CATGTTCAAAAGTGAAGACGGGTCTTGTCTGTCACCTTATAGATGATCGCGCTGATTA 254
QY 217 gctagcttgtaggttaagtcgcaaacgagcgagtcgtagcgagccttgagaag 276
DB 255 GCTAGTTGGTAGGTAACCGCTTACCAAGCAACGATACCTGACCCCTGAAGGGTGA 314
QY 277 ccgagcacatggaactgagacacggtccatactctacgaggaagcagcaatga 336
DB 315 TCGGCCACACTGGAAGTGAACACGGTCCAGACTCTTACGGAGAGGACGATGAGGAATC 374
QY 337 tgcgtcaatgaggaacacccctgaagcgaacgacgctgtaacgaagaagtcctga 396
DB 375 TTCCGCAATGGCGCAAGACCTGACGAGCAACGCCGCTGAGTAGTAGAGTCTCGGAT 434
QY 397 tgtaaagttcatltagcgaagaaaa-----taagcagaatgtagttagt 443
DB 435 CGTAAACTCTGTATTAGGGAAGACATATGTGTAAGTAACCTGTCCACATCTTGA 494
QY 444 acct--gcctaaagcagcgactaactagtcgaacgagcgcgtaatacgtatgca 501
DB 495 ACCTATTCAGAAAGCCAGCGGTAACTACGTGCGCAGCGCGGTATATAGTAGTGCA 554
QY 502 agcgctgttcggaatcattgagcgctaaaggctgctagcggatttgaagtagt 561
```

```
DB 555 ACCGTTATCCGGAATATTATGGCGTAAACGGCGGAGCGGCTTTTAAAGTCGATG 614
QY 562 aaaactgaggtctcaaccgctgctcgtcactgtaaacatacaagtcgttagag 621
DB 615 AAAGCCACAGGCTCAACCTGAGAGGCTCATTTGGAACCTGGAACCTTGACCAAG 674
QY 622 gcaatggaactccaggtgtagcgtgaaatgcgtagatatctgaggaacccaagtc 681
DB 675 GAAAGTGAATTCATGTTACGTAGCGGTAATATCGCAGACATATGGAAGAACAC 734
QY 682 gaagcgcactgctggtctaaacatgacgctgaggaacgaagcgtggttagtaac 741
DB 735 GAAGCGACTTCTGCTGTGTAAGTACGCTGATGTGCGAAGCGTGGGATCAAAAG 794
QY 742 attagatacccggtatccagccctaagctgtctaccagttgttgggggttt--a 799
DB 795 ATTAGTACCCTGTGATGTCACCGCTTAACGATGATGCTAAGTGTAGGGGTTTCCG 854
QY 800 accctcagtaacgaacctaagcgaattagacccgctggggaactatgctcgaagat 859
DB 855 CCCCTTAGTGTGCACTTAACGATTAAGCACTCCGCTGGGAGTAGACGCAAGCTT 914
QY 860 gaaactcaaggaatgacgaggtctcgcaacgaagcgtggaacatggtttaactgat 919
DB 915 GAAACTCAAAAGGAATGACGGGAGCCGCAACAGCGGTGAGCATGTGTTAATTGAA 974
QY 920 gataccccaacacccctcactggtgctgacatgtagatctaac-----atgtagagat 974
DB 975 GCAACCGCAAGAACCTTTACCAATCTTGACATCTTTGACAACTGAGATAGACCTT 1034
QY 975 atgagccctcgggagcagattcaacgaggtgcgaatggtgctgcgaagtcgtgc 1034
DB 1035 CCCCTTGGGGGACAAAGTGAAGAGTGTGATGTTGTCGTCAGCTGTCGTAGAT 1094
QY 1035 gtgggttaagtcocgcaacgagcgcaacccctatcgtatgtgcta-ccctaagttgg 1093
DB 1095 GTTGGGTAAAGTCCCGCAACGAGCGCAACCTTAACCTTAGTGGCCATTAAGTTGGG 1154
QY 1094 cactgttaacgaactgctcggttgacaacacggaggaagcggtgagtagtaactctca 1153
DB 1155 CACTTAACTTGACTGCGGTGACAAACCGAGAGAGTGGGATGACCTCAAAATCATCA 1214
QY 1154 tggccttlatgltccagggcacacacgctgctacaatgscggaatacagaaggtcgc 1213
DB 1215 TGGCCCTTATGATTTGGGTACACAGCTGTCAATGAGCAATFACAAAGGCGAAG 1274
QY 1214 cgcgaagaggagctaaactctctaaagtcggtccagttcgatgtgggtctgcaactcga 1273
DB 1275 CCGAGAGTCAAGCAAAATCCCATTAAGTTGTTCTCAATTCGAAATTGATCTCAACTCGA 1334
QY 1274 ccccatgaagctggaaatcgtctagtaacgagtaagcgtgcggtgtaactgtcc 1333
DB 1335 CTACATGAAGCTGGAATCGCTAGTATCTAGATGCACTGACGCTGAATACGTTCCC 1394
QY 1334 ggaactgtcacacacccgctgcacacacccctgagtgagggaacacccgaagtcgtct 1393
DB 1395 GGGTATTGTATACACCGCCCGCTACACACAGAGATTGTAACACCGCAAGCGGTGAG 1454
QY 1394 ttaacgtaaggaagacgaactactaagtgaaactcgttaaagggtggaagtcgttaaca 1453
DB 1455 TTAACCTTTAGAGAGTACCGCTGACAGGTGACCAAAATGATTGGGTGAAGTCTMAA 1514
QY 1454 ggtacc 1459
DB 1515 GGTAGC 1520
```

```
RESULT 7
US-09-198-955A-13
Sequence 13, Application US/09198955A
Patent No. 6187580
GENERAL INFORMATION:
```



```

; APPLICANT: Kim, Sang H
; TITLE OF INVENTION: Method for Producing Amide Compounds Using a Nitrite
; FILE REFERENCE: MSU 4.1-401
; CURRENT FILING DATE: 1999-02-10
; EARLIER APPLICATION NUMBER: 60/083,485
; EARLIER FILING DATE: 1998-04-29
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1517
; TYPE: DNA
; ORGANISM: Bacillus sp.
; FEATURE:
; NAME/KEY: rRNA
; LOCATION: (1)..(1517)
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: Z26929
; DATABASE ENTRY DATE: 1998-07-02
; US-09-248-528-2

```

```

Query Match      53.6%; Score 792.4; DB 3; Length 1517;
Best Local Similarity 75.0%; Pred. No. 4e-276;
Matches 1128; Conservative 0; Mismatches 326; Indels 50; Gaps 9;

```

```

QY 6 tggcttaagacacacgctgagcggtcttcaactgcaagtcgagcggtgtagcaata 65
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 tggcttaagacacacgctgagcggtcttcaactgcaagtcgagcggtgtagcaata 60
QY 66 cc-----taagcgcaacggtgtagcaacgctgtagcttccctccg-ag 110
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 61 gctgtcctcttaagtgtagcgagcgagcggtgtagcaacgctgtagcaacgctgtagc 120
QY 111 tctggataacttccgcaagaagataaccgagtagctctgttgatacaagaatt 170
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 121 actggataacttcggaacacggaagataaccggaacacggaacacggaatt 180
QY 171 tgataagtaag-----attatctgttgagatgagccgagcgagattagtagt 222
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 181 tgggttaagagcgagctttagctgtcagcaggaatggcgccgagcagattagtagt 240
QY 223 tggtagagtaatgctcaccacgaagcgagtagcggcctgtagaggtgtccggcc 282
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 241 tggtagagtaatgctcaccacgaagcgagtagcggcctgtagaggtgtccggcc 300
QY 283 acaatggaactgagacaggttccatctccctacgagagcgagcagattaaagaattctgtctc 342
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 301 acactggagactgagacagcgccagactcctacgagagcgagcagtaggaattctccgc 360
QY 343 aatgggggaacacctgaagcagcgagcggtggaacggaaggtcttcgagtttaaa 402
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 361 aatggagcaagctgagcaggaacacgagcggtggaaggaaggtcttcgagtttaaa 420
QY 403 gtccatagcaggaanaataaagcagcaatgtga-----tgatgttacctg 448
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 421 gctctgttgtaggaagaacagtagcgttgcgaacagggcggttacctgtagcgttacctg 480
QY 449 ccta--aagcagcgcttaactacgtgcagcagcgcggttaactgtagtgcgaagcgt 506
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 481 acgaagaacacgaagctacgtacgtgcagcagcgcggttaactgtagtgcgaagcgt 540
QY 507 tgttcggaatcatctggcgcttaaggggtgcgtagcgagattgtgaagtgcgaagc 566
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 541 tgcgcgaatattctggcgcttaaggggtgcgtagcgagattgtgaagtgcgaagc 600
QY 567 tgcgcgctcaacccggtgcgtgacttgaactgaactgaactgtgaggttgggaagcgcaag 626
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 601 tgcgcgctcaacccggtgcgtgacttgaactgtgaggttgggaagcgcaaggtgaggtgag 660
QY 627 tggaaatccaggtgtgtagcggtggaatgtagatcttcgaggaacacacaggtgcgaag 686
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

DB 661 cggaaatccacgttagcggtgaatgtcg tagagatgttggaagaaacacacagctgagcgaa 720
QY 687 cgactgtgctgctcaaaactgacgtgaggaacggaagcggtgtagtaaacggagtag 746
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 721 cggctctctgagcctgttaactgaacgtgagcggaagcggtgtagtaaacggagtag 780
QY 747 ataaccggttaatccacgcccctaaacgtgtgtctacaggtgtgttggtgtttt--aacct 804
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 781 ataaccgtgttagtccacgcccgttaacgaatgtagtgcgaagtgttagaggttagtccacccct 840
QY 805 cagtaagaacactaagcagatlaagtagacgccttgggagctatgtctgcgaagaatgaaac 864
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 841 tagtgcgtgcaagaacgacgtacgtccgcctggggagtagtcgagcggaagcgtgaaac 900
QY 865 tcaaaagaaatgacgggggttccgcacaacggttggagcgtgtgttaattgtagatgc 924
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 901 tcaaaagaaatgacgggggttccgcacaacggttggagcgtgtgttaattgtagatgc 960
QY 925 cccaaaacccctcagcctgagcttgaacat-----gagatgatacatgtagagatata 977
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 961 ggaagaacacttaacacagctgtgacatcccccagacaacccctagagtagagcggtccct 1020
QY 978 agccttcgggcaagatcacaggtgctgcatgtgttcgtcaagctcgtgtgagatgt 1037
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1021 ttcgggggagacaggggtgagcaggtgtgtcgtcagcagctcgtgtgtgagatgt 1080
QY 1038 ggggttaagtcggcgaacgagcggaacccctatcgtgtgtgtgcta--ccttaagtgtggac 1096
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1081 ggggttaagtcggcgaacgagcggaacccctatcgtgtgtgtgcta--ccttaagtgtggac 1140
QY 1097 tggtagaactgctcggttgacaaaacggaaggaagcgggagtagacgttcaatctcattg 1156
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1141 tctaaagtgactcgcgctaaagaatcggaaggtggaaggtggaagtagcgtcaatcatatgc 1200
QY 1157 ccttatgtccagggccacacacgctgtctacaatggtcggtacaggaaggtgcacacatgc 1216
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1201 cccctatgacccctgtgtacacacgctgtcaaatggtgtgtaaaaaggaagcgaaacccgc 1260
QY 1217 aagaggaactaactctcctaaagtctgtccagcttgcgagttgggtgtcgaactgaaccc 1276
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1261 gagggtgagcgaatcccaaaaacacactctcagttcggattggaagctgtgaactgtccgc 1320
QY 1277 catgaagctggaatcgtctagtaactgcgagatcagcatctgcgcggtgataactgtccgga 1336
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1321 catgaagcggaaatcgtctagtaactgcgagatcagcatctgcgcggtgataactgtccgga 1380
QY 1337 cctgttacaacacgcccgttcaacacacactgaggtggggagacacccggaagtgtcttgtta 1396
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1381 tctgttacaacacgcccgttcaacacacactgaggtggggagacacccggaagtgtcttgtta 1440
QY 1397 accgtta--aggaagacagactaaggtgaacactcgtlaaaaggggtgaaatcgttaacaaag 1455
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1441 ccccttcgggagcgaacccgcggaaggttggaacaaatgtagtgggtgaaatcgttaacaaag 1500
QY 1456 tacc 1459
    |||
DB 1501 tagc 1504

```

```

RESULT 9
US-09-549-108-2
; Sequence 2, Application US/09549108
; Patent No. 6214603
; GENERAL INFORMATION:
; APPLICANT: Oriol, Patrick J
; APPLICANT: Padmakumar, Rugmini
; APPLICANT: Kim, Sang H
; TITLE OF INVENTION: Method for Producing Amide Compounds Using a Nitrite
; FILE REFERENCE: MSU 4.1-486
; CURRENT APPLICATION NUMBER: US/09/549,108
; EARLIER FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/083,485

```

```
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 09/248,528
; PRIOR FILING DATE: 1999-02-10
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1517
; TYPE: DNA
; ORGANISM: Bacillus sp.
; FEATURE:
; NAME/KEY: rRNA
; LOCATION: (1)..(1517)
; DATABASE ACCESSION NUMBER: Z26929
; DATABASE ENTRY DATE: 1998-07-02
; US-09-549-108-2
```

```
Query Match      53.6% Score 792.4; DB 4; Length 1517;
Best Local Similarity 75.0%; Pred. No. 4e-276;
Matches 1128; Conservative 0; Mismatches 326; Indels 50; Gaps 9;
```

```
QY 6 tggctcagaactaacgctgctgcgcgcgtcttaacaatgcaagtcgagcggtgtagcata 65
Db 1 tggctcagaactaacgctgctgcgcgcgtcttaacaatgcaagtcgagcggtgtagcata 60
QY 66 c-----tagcgcgcaacgggtgtagtaacagctgtgtaattcttcctcg-ag 110
Db 61 gctgtctccttaagcttagctgagcgagcggtgtagtaacagctgtgtaacagctgtcgag 120
QY 111 tctggataactcttcgaaagaagtaataacgagatgctcgtgttgatcataaagatt 170
Db 121 actggataactcttcgaaagaagtaataacgagatgctcgtgttgatcataaagatt 180
QY 171 tgataagtaag-----attattgcttgagagatgagcccgcgccgatttagctagt 222
Db 181 tggctgagaagcgagcttcttagctgtcactgcaagatgagcccgcgccgatttagctagt 240
QY 223 tgggagagtaattgcttcacaaagcgagatcggttagcgccgctgagaggggtgtccggcc 282
Db 241 tgggagagtaattgcttcacaaagcgagatcggttagcgccgctgagaggggtgtccggcc 300
QY 283 acaatgagacgtgagacagctgtcactactcctacaggaagcgagtaagaatttgctc 342
Db 301 acaatgagacgtgagacagctgtcactactcctacaggaagcgagtaagaatttgctc 360
QY 343 aatgggggaaacctggaagcgagcgccggttgaacgaagaaggtcttcgattgtaaa 402
Db 361 aatgggggaaacctggaagcgagcgccggttgaacgaagaaggtcttcgattgtaaa 420
QY 403 gttaactagcgagaaataaagcagcaattgta-----tgaatgtacctg 448
Db 421 gctcgtgtgtaaggaagaagaagtaacgcttcgaaacagggcggtaccttgtaacgtaactg 480
QY 449 ccta--aagcagcggtactactgctgacagcagcgcggttaactgtaatgctgtaacgct 506
Db 481 acgaaagcgagcagcgtactactgctgacagcagcgcggttaactgtaatgctgtaacgct 540
QY 507 tgttgagatacttgagcgtaaggttgctgagcgagatttgtaagtaaggtgtaaac 566
Db 541 tgttgagatacttgagcgtaaggttgctgagcgagatttgtaagtaaggtgtaaac 600
QY 567 tggcgagctcaacccgtggtccttgacttgaaactaactgtaagcttggttggaagagcaag 626
Db 601 tggcgagctcaacccgtggtccttgacttgaaactaactgtaagcttggttggaagagcaag 660
QY 627 tggaaatccaaagtgtagcggtggaatgctgatatcgtggaagaaacacagctggtcgaaag 686
Db 661 cgaattccaaagtgtagcggtggaatgctgatatcgtggaagaaacacagctggtcgaaag 720
QY 687 cgaattcgtggtcgaacaaactgagcggtggaagcgagaaaggtggtgtagtaaacgggattag 746
Db 721 cggctctcgtggtcgaacaaactgagcggtggaagcgagaaaggtggtgtagtaaacgggattag 780
```

```
QY 747 ataccggtaactccagcccttaacgctgtctaccagttgttgagggttt--aaccc 804
Db 781 ataccggtaactccagcccttaacgctgtctaccagttgttgagggttt--aaccc 840
QY 805 caatgaacactcaacgtaagtaagtagaacgcttggtggaactatgctgcgaagtgaaac 864
Db 841 tagtgcgagcaaacgcatlaagcaactccgctggtggtgtagcgcgcgaagctgaaac 900
QY 865 tcaaaagaaatgacgggtgtccgcacaagagcggtggtgagcattgttaattgtagatag 924
Db 901 tcaaaagaaatgacgggtgtccgcacaagagcggtggtgagcattgttaattgtagatag 960
QY 925 ccaaaaacactcccttgctgactacat-----ggatctgaaatcagtagagataatg 977
Db 961 gcaagaaacttaacaggtgtgacatcccttgacaacacttagagatagggtctccct 1020
QY 978 agccttcggagatccacaggtgtgcatggtgtgtcgttaagctgtgtcgtgtagatgtt 1037
Db 1021 ttcgggggagacaggtgtagcaggtgtgtgcatggtgtgtcgttaagctgtgtcgtgtagatgtt 1080
QY 1038 ggtgttaagctccgcgaacagcgacacccctatcgtatgtgtcta-ccctaagttggtgac 1096
Db 1081 ggtgttaagctccgcgaacagcgacacccctatcgtatgtgtcta-ccctaagttggtgac 1140
QY 1097 tggtaacgaactgctgcgttgacaacacggaggaagcggtggaatgagctcaaatccctatg 1156
Db 1141 tctaagtgtagctgcgtgttaaaagtgcgaggaaggtggtgtagtgcataatcatactgc 1200
QY 1157 ccttaatgtccagggccacacacgtgtctataatggtccgataacagaggtgtgccaactgc 1216
Db 1201 ccttaatgtccagggccacacacgtgtctataatggtgtgtaacaaaggcgagcaaacgc 1260
QY 1217 aagagggagcttaactctctaaagtcggttccaggtcgtgattggtgtgaactgcagcc 1276
Db 1261 gagggtgagaggaatcccaaaaacactctcagttcgtgattggtgaggtgtgaactgcagctg 1320
QY 1277 catgaagtgcgaatcgtcgtgtaatcgtgcgagatcagcatgctgcggtgtgaatacgttccgga 1336
Db 1321 catgaagcgggaatcgtcgtgtaatcgtgcgagatcagcatgctgcggtgtgaatacgttccg 1380
QY 1337 cctgtgaacaccccgctacacacactggtgtggtgagcccggaaggtgtcttcttga 1396
Db 1381 tctgtgaacaccccgctacacacactggtgtggtgagcccggaaggtgtcttcttga 1440
QY 1397 acgcta-aggagagactactaaggtgaaactcgtlaaagggttgaagtgttaacaag 1455
Db 1441 ccttaacgggagcagccgcggaaggtggtgacaaatgtatgtggtggaagtgttaacaag 1500
QY 1456 tacc 1459
Db 1501 tagc 1504

RESULT 10
US-09-549-111-2
; Sequence 2, Application US/09549111
; Patent No. 6228633
; GENERAL INFORMATION:
; APPLICANT: Oriol, Patrick J
; APPLICANT: Padmakumar, Rugmini
; APPLICANT: Kim, Sang H
; TITLE OF INVENTION: Method for Producing Amide Compounds Using a Nitrile
; FILE REFERENCE: MSU 4.1-489
; CURRENT APPLICATION NUMBER: US/09/549,111
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/083,485
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 09/248,528
; PRIOR FILING DATE: 1999-02-10
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
```

Accession	Sequence	Length	Source
Db 841	tagtctcgcagcgaacgcattaaagcattccgccttgaggagtaagcgcgaagcgcgaac	900	US-09-549-106-2
QY 865	tcaaaaggaattgacgggggtccgcacaagcggctggagacatggtttaaattcgatgatac	924	Sequence 2, Application US/09549106
Db 901	tcaaaaggaattgacgggggtccgcacaagcggctggagacatggtttaaattcgatgatac	960	Patent No. 6242242
QY 925	cccaaaacctactctggttgatc-----gatctgaattcattgaagataatctg	977	GENERAL INFORMATION:
Db 961	gggaagaaccttaccaggatcttgacatcccttcgacaacccctagaagataggcgttccct	1020	APPLICANT: Oriol, Patrick J
QY 978	agccttcggagcagatctcaacagtgctgatatgctgctcaagctcgtatcgttgagatct	1037	APPLICANT: Padmakumar, Rugmini
Db 1021	ttcgggggggacaggggtgacaggttggtgacatgctgctcaagctcgttctggagatct	1080	TITLE OF INVENTION: Method for Producing Amide Compounds Using a
QY 1038	gggttaagttccgcacaagcgaagcgaacccctacatcgtatgtgcta-ccttaagttggcac	1096	FILE REFERENCE: MSU 4.1-487
Db 1081	gggttaagttccgcacaagcgaagcgaacccctacatcgtatgtgccaagcttcaagttggcac	1140	CURRENT APPLICATION NUMBER: US/09/549,106
QY 1097	tgtatcgaaactgcgcgttgacaaacccggagagaagcggggatgtagcgttaaatctctatgg	1156	PRIOR FILING DATE: 2000-04-13
Db 1141	tctaaagtgtactgcggtctaaagaatcggagagaagctggggatgtagcgttaaatctctatgg	1200	PRIOR APPLICATION NUMBER: 60/083,485
QY 1157	cccttaattgccaggccacacacgctgctacaaatggccgatacagaaggggtcccaactgcg	1216	PRIOR FILING DATE: 1998-04-29
Db 1201	cccttaattgccaggccacacacgctgctacaaatgggtggttaaaaaggccagcgaacccgc	1260	PRIOR APPLICATION NUMBER: 09/248,528
QY 1217	aaggaggaagctaattcttaaaagtcggtccacgttcggatltgggtctcgcgaactcgacc	1276	NUMBER OF SEQ ID NOS: 18
Db 1261	gggggtggagcgaattcccaaaaaaacacactctgaatctcgatttgcaagctcgcaactgcgt	1320	SOFTWARE: Patenlin Ver. 2.0
QY 1277	catgaagtcggaatctcgatgtaatcgcggatcagcagtcgacgcgcgttgaaatcgtccgga	1336	SEQ ID NO 2
Db 1321	catgaagtcggaatctcgatgtaatcgcggatcagcagtcgacgcgcgttgaaatcgtccgga	1380	LENGTH: 1517
QY 1337	ccctgtacacacggccgcgttcacacacccctgagctgggggacacccgaatgtgcttctgta	1396	TYPE: DNA
Db 1381	tctgttaacacccgcgcgttcacacacagagatttgttaacacccgaatgtgcttctgta	1440	ORGANISM: Bacillus sp.
QY 1397	accgta-aagaaacagcgaactcaaggtgaactcgttaaaagggtgggaagtcgtacaag	1455	FEATURE:
Db 1441	cccttaacgggaacccgcgcgcgaaggtggacaatgatttgggtcgtaagtcgtacaag	1500	NAME/KEY: rRNA
QY 1456	tacc 1459		LOCATION: (1)...(1517)
Db 1501	tagc 1504		

DATABASE ACCESSION NUMBER: 226929
 DATABASE ENTRY DATE: 1998-07-02
 US-09-549-106-2

Query Match 53.6%; Score 792.4; DB 4; Length 1517;
 Best Local Similarity 75.0%; Pred. No. 4e-276;
 Matches 1128; Conservative 0; Mismatches 326; Indels 50; Gaps 9;

```

QY 6 tggctcaagaactaacgttgcgagcgctcttaacaatgcagtcgagcgagtagcaata 65
Db 1 tggctcaagaactaacgttgcgagcgctcttaacaatgcagtcgagcgagtagcaata 60
QY 66 c-----tagcgcgcaacggttgagtaacacgttgtaactctcctccg-ag 110
Db 61 gcttgctccttaagtttagcgcgagcggttgagtaacacgttgagcaactgccttcgag 120
QY 111 tctgggaatacttcggaaggaagtaataccgagtagtctcgttgatcaacaagatt 170
Db 121 actgggaatacttcggaaggaagtaataccgagtagtctcgttgatcaacaagatt 180
QY 171 tgaataagtaag-----attatctgtgagatgagcccgcgagcagattagtaagt 222
Db 181 tctggttgaagaagcgctcttaagctgttcaactgcagagatggcgccgagcagattagtaagt 240
QY 223 tggtagagtaatgagtcacaagaagcgagatcggttaacgagcctgagagaggtgcgggccc 282
Db 241 tggtagagtaatgagtcacaagaagcgagatcggttaacgagcctgagagaggtgcgggccc 300
QY 283 acaatgtgaactgagacacgagtcacatactcctacgagagcgagcagattaaagatctgtc 342
Db 301 acaatgtgagacgagacacgagtcacatactcctacgagagcgagcagattaaagatctgtc 360
QY 343 aatgggggaaacccctgaaagcagcgagcggtgaaacgaagaaggtcttcgagattgtaaa 402
Db 361 aatggagcaaatcttgagcagcgagcgagcggtgaaagaggtcttcgagattgtaaa 420
QY 403 gttcatctgagcagaagaataagtaagcagaattgta-----tgaatgttacctg 448
Db 421 gttcatctgagcagaagaataagtaagtaagcagaattgta-----tgaatgttacctg 480
QY 449 ccta--aagcacccgagctactactgagcagcagcgcgtaatacgtatgagtaacagcgt 506
Db 481 acgaggaagacacgagctactactgagcagcagcgcgtaatacgtatgagtaacagcgt 540
QY 507 tgttcggaatacttgagcgtaaaaggtgagcgagaggttgtaagtaaggtgtaaaac 566
Db 541 tgttcggaatacttgagcgtaaaaggtgagcgagaggttgtaagtaaggtgtaaaac 600
QY 567 tgcgggctcaaacccgttgagccttgcaacttgaaactgaaacttgagagagagcaag 626
Db 601 tgcgggctcaaacccgttgagccttgcaacttgaaacttgagagagagagagag 660
QY 627 tggaaatcccaagtgtagcggtgaaatgtagatctcggaggaacacacagtgagcgagag 686
Db 661 cggaaatcccaagtgtagcggtgaaatgtagatctcggaggaacacacagtgagcgagag 720
QY 687 cgaacttgagcggttcaaaactgagcgagcaggaagcggttgagtagtaaaagagagtag 746
Db 721 cggacttcgagcggttcaaaactgagcgagcaggaagcggttgagtagtaaaagagagtag 780
QY 747 ataaccggttaataccacgcttaaacggttgtagcaagttgttgagggtttt--aacct 804
Db 781 ataaccgttgtagtccagcgagcgttaaacgagtgtagtgaaggtgtatccacctt 840
QY 805 cagtaacgaactaacgagtaagtagagacgagcgtgagagtagtctcggaggaagagtaaac 864
Db 841 tagtgcgtgagcaaaacgagcagctcgcgttgagagtagtgcgagcgagagcgagtaaac 900
QY 865 tcaaaaggaattacgaggggttcgcgacaagcggtgagcagtgtagtttaattcgaatatag 924
Db 901 tcaaaaggaattacgaggggagccgcgacaaagcggtgagcagtgtagtttaattcgaataaac 960

```

```

QY 925 cccaaaacactcaaccgtggcttgagac-----gagtcgatacatgtagagatata 977
Db 961 gcgaaagaaccttaccacgggtcttgagacatccctcgacaacacctgagataggggttccct 1020
QY 978 agccttcggagcagatcatcaaggtgtagcagtagtgcgtcagcgtgctgtagagttt 1037
Db 1021 ttcggggggagcaaggtgtagcagtagtgcgtcagcgtgctgtagagttt 1080
QY 1038 gggtagtcccgcaacgagcgcaacaccttalcgtatgtagtga--ccttaagttagggc 1096
Db 1081 gggtagtcccgcaacgagcgcaacaccttalcgtatgtagtga--ccttaagttagggc 1140
QY 1097 tggtagaacttcggtgtagcaaacggaggaagcgaggtgagtagcgtcaaatcctcatgg 1156
Db 1141 tctaaagttagctgcggctcaaaaagtctgaggaaggtgaggtgagtagcgtcaaatcctcatgg 1200
QY 1157 ccttatgtccagggccacaacacgtgtagcaaatgtagcagatagagaggtgcgcaactgcg 1216
Db 1201 ccttatgtacgttgtagtacaacacgtgtagcaaatgtagtgcagaggtgcgcaactgcg 1260
QY 1217 aagagggagtagtctcttaaaagtccggtccaggttcgagattgaggtcgaactcgacc 1276
Db 1261 gaggtagaggaatcccaaaaacactcaggttcgagattgaggtcgaactcgacc 1320
QY 1277 catgaagtcggaatcgtagtaatcgtagtcagcagatgcgagtagaatacgttcccgga 1336
Db 1321 catgaagtcggaatcgtagtaatcgtagtcagcagatgcgagtagaatacgttcccgga 1380
QY 1337 cctgttacaacacgcccgttaccacacctgaggtgggagagcccgaggtggtcttgta 1396
Db 1381 tctgttacaacacgcccgttaccacacctgaggtgggagagcccgaggtggtcttgta 1440
QY 1397 acgcta-aggaagacactactaaggtgaactcgttaaaaggggttgaagtcgttaacaag 1455
Db 1441 cctttagaggaagcagcagcgcgaggaaggtgaggaacaatgattgggtgagtcgttaacaag 1500
QY 1456 taacc 1459
Db 1501 tagc 1504

```

```

RESULT 12
US-09-550-394-2
; Sequence 2, Application US/09550394
; Patent No. 6287828
; GENERAL INFORMATION:
; APPLICANT: Ortel, Patrick J
; APPLICANT: Padmakumar, Rugmini
; APPLICANT: Kim, Sang H
; TITLE OF INVENTION: Method for Producing Amide Compounds Using a Nitrite
; FILE REFERENCE: MSU 4.1-488
; CURRENT APPLICATION NUMBER: US/09/550.394
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/083.485
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 09/248.528
; PRIOR FILING DATE: 1999-02-10
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1517
; TYPE: DNA
; ORGANISM: Bacillus sp.
; FEATURE:
; NAME/KEY: rRNA
; LOCATION: (1)..(1517)
; DATABASE ACCESSION NUMBER: 226929
; DATABASE ENTRY DATE: 1998-07-02
; US-09-550-394-2

```

Query Match 53.6%; Score 792.4; DB 4; Length 1517;

Best Local Similarity 75.0%; Pred. No. 4e-276;
Matches 1128; Conservative 0; Mismatches 326; Indels 50; Gaps 9;

```
OY 6 tggctaagaactaacgctgacgagcgtcttaacaatgcgaatcgcgagcggtagacaata 65
    ||||||| || ||||||| || || ||||||| || || ||||||| || || ||||||| ||
Db 1 tggctaagaactaacgctgacgagcgtcttaacaatgcgaatcgcgagcggtagacaata 60
OY 66 cc-----taagcgcaacgggtgagtaacacgttgtaactctcccg--ag 110
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 61 gcttgcctcttgatgtagcgacgagcggtgagtaacacgttgtaactctcccg--ag 120
OY 111 tctggataactctccgaaaggaaagtaaacggaatagcttctgtgtgtaacaaagt 170
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 121 acgggataactctccgaaacggagcttaacacggaataacacggaacacggaatgttt 180
OY 171 tgaatgtaag-----attatgtcttgagatgagccgagcggaatagtagt 222
    || || || || || || || || || || || || || || || || || || || || || ||
Db 181 tggatgtaagcgagctttagctgtacatgcagagatgagccgagcggaatagtagt 240
OY 223 tggatgtaagctgacacaaaggcgagatcggtagccgagcttgagagaggtctccggc 282
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 241 tggatgtaagctgacacaaaggcgagatcggtagccgagcttgagagaggtctccggc 300
OY 283 acaatgtaactgagacacggttccatactcctacgagagcgagcagatgaatctgtctc 342
    || || || || || || || || || || || || || || || || || || || || || ||
Db 301 acaatgtaactgagacacggttccatactcctacgagagcgagcagatgaatctctcgc 360
OY 343 aatgggggaaaccttgaaagcagacgagcggtgaaagaaagaggtcttctggatgtaaa 402
    || || || || || || || || || || || || || || || || || || || || || ||
Db 361 aatgggaaagcttgcagagacacggtgaaagaaagaggtcttctggatgtaaa 420
OY 403 gttcatgagcagaaataaagcaagcaatgta-----tgatgtaactg 448
    || || || || || || || || || || || || || || || || || || || || || ||
Db 421 gttcatgagcagaaataaagcaagcaatgta-----tgatgtaactg 480
OY 449 ccta--aagcaccggttaactacgtgacgagcgcggtgaatacgtatggtgcgaagcgt 506
    || || || || || || || || || || || || || || || || || || || || || ||
Db 481 acgaaagaaagcagacggttaactacgtgacgagcgcggtgaatacgtatggtgcgaagcgt 540
OY 507 tgttggaaatcatctggcgtaaaaggggtgagtagcgatcttgtaagtcagatgtgaaac 566
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 541 tgttggaaatcatctggcgtaaaaggggtgagtagcgatcttgtaagtcagatgtgaaac 600
OY 567 tggcggtcacaacccgtgacgtcactgtaaacatacaagcttgagatcttggaagagcaag 626
    ||||||| || || || || || || || || || || || || || || || || || || || ||
Db 601 tggcggtcacaacccgtgacgtcactgtaaacatacaagcttgagatcttggaagagagag 660
OY 627 tggaaatccaggtgtagcggtgaaatgctgataatcttggaagaaacacacggtgcgaag 686
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 661 cggaaatccacggtgtagcggtgaaatgctgataatcttggaagaaacacacggtgcgaag 720
OY 687 cgaatgctggtcacaacacgagcgtgagacgaagaagcgttggtgtagtaacgggtag 746
    || || || || || || || || || || || || || || || || || || || || || ||
Db 721 cggatcctctggtcacaacacgagcgtgagacgaagaagcgttggtgtagtaacgggtag 780
OY 747 ataccgcggtgaatccacacgaccttaacgctgtctacacgagtggtgggggttt--aacct 804
    ||||||| ||||||| ||||||| ||||||| || || || || || || || || || || || ||
Db 781 ataccgcggtgaatccacacgaccttaacgagtagtgtagtaaggttagagatccacctt 840
OY 805 cagtaacgaactcaacgagtagtagacgagcgttggaactatgctgcgaagatgaaac 864
    || || || || || || || || || || || || || || || || || || || || || ||
Db 841 tagtgctgcagcaacacgagtagtagacacgctcggtggaatgtagcgcgaagcgtgaaac 900
OY 865 tcaagaagaatcgagcggtggtccgacaaagggtggaagcagtggttaattcatgtagt 924
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 901 tcaagaagaatcgagcggtggtccgacaaagggtggaagcagtggttaattcatgtagt 960
OY 925 cccaataacacacgagcgttgtagt-----ggaatcgaatcagtagtagatag 977
    || || || || || || || || || || || || || || || || || || || || || ||
Db 961 gcgaaagaaacttaacaggtcttgacatcccttgacaaacaccttagagatagggcgttccct 1020
OY 978 agccttcgcaagatcattacaggtgctgcagatggtgtgtcgtacagctggtcgtgagatgtt 1037
    || || || || || || || || || || || || || || || || || || || || || ||
```

```
Db 1021 ttcgggggacaaagggtgacaggtggtgcatggtgtgtcgtacgctgtgtgtgagatgt 1080
OY 1038 ggtttagtcccgcaacgaagcgcaacccctatgctgtgtgcta--ccttaagtgtggac 1096
    ||||||| ||||||| ||||||| ||||||| || || || || || || || || || || || ||
Db 1081 ggtttagtcccgcaacgaagcgcaacccctatgctgtgtgcta--ccttaagtgtggac 1140
OY 1097 tggtaacaaactgcggtgtaacaaacgggaagcggtgagtagcgttcaatctcatg 1156
    || || || || || || || || || || || || || || || || || || || || || ||
Db 1141 tctaaagttagctcggcttaaaagtcgggaaggtggtggaatgtagcgttcaatctcatg 1200
OY 1157 cctttagtccagggccacacacgctgtacaaatggtccgaatacagaggttcgaactgc 1216
    || || || || || || || || || || || || || || || || || || || || || ||
Db 1201 cctttagtccagggccacacacgctgtacaaatggtccgaatacagaggttcgaactgc 1260
OY 1217 aagaagagctaaatctctaaagtgtcccaatggtgagatggtggtgtcgtgaactgaccc 1276
    || || || || || || || || || || || || || || || || || || || || || ||
Db 1261 gagggtgagcgaatcccaaaaacacacatcctcaatggtgagatggtggtgagcgtgac 1320
OY 1277 catgaatcgggaatcgtgtaataatcgtggaatcgaatgcggtgatacgttccgga 1336
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1321 catgaatcgggaatcgtgtaataatcgtggaatcgaatgcggtgatacgttccgga 1380
OY 1337 cctttagtccagggccacacacgctgtacaaatggtccgaatacagaggttcgtgtta 1396
    ||||||| ||||||| ||||||| || || || || || || || || || || || || || ||
Db 1381 tctttagtccagggccacacacgctgtacaaatggttcgaatcgtggtgtgtgtgt 1440
OY 1397 accgta--aggaagcagactcctaaggtgaactcgttaaaagggtgtaagtcgttaacaa 1455
    || || || || || || || || || || || || || || || || || || || || || ||
Db 1441 cctttagtccagggccacacacgctgtacaaatggttcgaatcgtggtgtgtgtgtgt 1500
OY 1456 taac 1459
    || ||
Db 1501 tagc 1504
```

```
RESULT 13
US-09-193-377B-3
; Sequence 3, Application US/09193377B
; Patent No. 6221594
; GENERAL INFORMATION:
; APPLICANT: Burrell, Paul
; APPLICANT: Blackall, Linda
; APPLICANT: Keller, Jurq
; TITLE OF INVENTION: METHOD FOR THE DETECTION OF AQUATIC
; TITLE OF INVENTION: NITRITE OXIDISING MICROORGANISMS OF THE GENUS NITROSPIRA
; FILE REFERENCE: CULINDO 001AUS
; CURRENT APPLICATION NUMBER: US/09/193, 377B
; CURRENT FILING DATE: 1998-11-17
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1500
; TYPE: DNA
; ORGANISM: Nitrospira
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(1500)
; OTHER INFORMATION: n - A,T,C or G
US-09-193-377B-3
```

Query Match 53.5%; Score 790.6; DB 4; Length 1500;
Best Local Similarity 74.8%; Pred. No. 1.8e-275;
Matches 1121; Conservative 0; Mismatches 341; Indels 37; Gaps 9;

```
OY 1 gatcatggtcagaactaacgctggtggtggtcgtcttaacaatgcgaatcgaagcggt---- 56
    ||||||| ||||||| ||||||| ||||||| || || || || || || || || || || || ||
Db 3 gatcctggtcagaactaacgctggtggtggtcgtcttaacaatgcgaatcgaagcggtgaga 62
OY 57 -gtagcaatacct-----agcgcggaagcggtgagtaaacagt--ggttaattctcccg 108
    ||||||| || || || || || || || || || || || || || || || || || || || ||
Db 63 cgtagaatacgttctgtaaacggtgagcggtgagtaaacatcacatggttaacacacccctg 122
```

109 agtctggataacttccgaaagaaagcctaataccggaatagctcttggatacaaga 168
123 agtggggaataactagcgaagaaagtgactaataccgatactcctgctcggat 182
169 ttgtatggataaagc-----tttatgtctggagataagcccgccggaattcg 219
183 cggagagaaagcgaataaccggtggtatccgctcttgatgctgctcaatgctcaagct 242
220 agtctgtgaggttaatgctcaacaaagcgcagatcgtatagccgctgagaaggtgtccg 279
243 tgtgtgtgaggttaacggtctcaacaaagcgtctgacggtgtgtgtgagagagcgaatca 302
280 gcccaatggagacgtgagacgtgcatcactcaacgaggaagcgaactgaagatctgt 339
303 gcccaactgagacgtgagacgtgagacgtcctcaagggagagcgaactgaagatctgt 362
340 ctcaatggagaaacccctgaaagcagcgcgcgtgaaacgaagaaggtctcgatgt 399
363 cgcaatggcgcagacgtcgcagc 422
400 aaagtctatggcagagaaat---aagcagcaatgtgagatgctcctgctcaaa--a 454
423 aaaccccttcgycgagggaaagatgaacggttaacccgttcggaacgtctcgcaagaagca 482
445 gcccggttaactagctgcagcagcgcgcgttaatactatgctgcaagcgttctcga 514
483 gccacggttaactctgcgcagcgcgcgcgcgttaatactatgcaagcgttctcga 542
515 atcaatggcgttaaaaggtgcgtgagcgaattgtgaatcaggtgtgaaactcgcgcgc 574
543 tttaactggcgttaacagggagcgtgagcgttggtaagccctcctgaaatctccggcgc 602
575 caacccggtgcgttcgaactgaacatacaggttggaggttggaggaagcgaattgaatc 634
603 taacccggaagagtgccgaggggagcgcgcgcgcgttaagatgtggaagagagcgcgaatc 662
635 caggtgtgagcgttgaatgtgagatctgagaggaacacacagtcgcaagcgcgcgcgc 694
663 ccggtgtgagcgttgaatgtgagatctgagaggaagcgcgcgcgcgcgcgcgcgcgcgcgc 722
695 tggcctcaaacacgacgtcgaagcagcaagagcgtggtgagtaacggggttaataccgcgc 754
723 tggacacattctcgacgctgaggtctcgaagcgttggggaacacacggggtatgataccgcgc 782
755 gaaatccacgccttaacgctgtctacaggttctgaggggtttaaaccctcgaatgaagaa 814
783 gtagtccacgccttaacgctgtctacaggttctgaggggtttaaaccctcgaatgaagaa 840
815 cctaacggaatgaatgaacgcgcgttggggaactgtctcgaagagtgaaactcaagaagaa 874
841 gctaacgcataatgaataccgcgcgttggaggtagtcgcgaaggttgaactcaagaagaa 900
875 tgaacggggttcgcgaacacggttggagcgttgggttaattcgaatgatacccaaaac 934
901 tgaacggggttcgcgaacacggttggagcgttgggttaattcgaatgatacccaaaac 960
935 tcaacggtgagcgtgaacgctgaatcgtgaatcgtgaatgatacccaaaacggttgcgcgcgc 991
961 ttaacccagcgtgacacgtgaggtgagtaagaggtgaaagccttaacgaagtgaaacacaa 1020
992 ---ttcaacggtgacgtgaggtgaggtgaggtgaggtgaggtgaggtgaggtgaggtgag 1048
1021 tctctgtcgtgaggtgaggtgaggtgaggtgaggtgaggtgaggtgaggtgaggtgaggtgag 1080
1049 cgaacgagcgaacacccctcgtatgtctgacac---ttaagtgtgagcagctgatacga 1105
1081 cgaacgagcgaacacccctcgtatgtctgacacacacggttcaagcgcgggaactctgagag 1140
1106 actgctgagcgtgaacacgagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcga 1165
1141 actgctgagcgtgaacacgagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcga 1200
1166 cgaagcgaacacgctgctacaaatgctcgaatacagaggtgtcgcgaactcgcgaagagagag 1225

1201 ctggggccacacacgctctcaaatggtccggttaacaaagcgtctgcaaacccgtaaggggag 1260
1226 ctaactctaaagtggtccagctcgatggtgtggtctgcaactcgaacccatgaagtc 1285
1261 ccaatcgcaaaacacggccctcagttcagatgtgagcttcaactcgaactcgaatgaagtc 1320
1286 ggaatcgttaatgctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1345
1321 ggaatcgttaatgctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1380
1346 caccgcgcgtacacacacgtgagtgaggagacacccgaagtggtcttctgaacgtaag 1405
1381 caccgcgcgtacacacacgtgagtgaggagacacccgaagtggtcttctgaacgtaag 1440
1406 agaacgacttaagtggaactcgttaaggggtggaagtcgtaacgaagtaacgtaac 1464
1441 gggcagacgccacagatagcagcagatgtggtgaggtgaggtgaggtgaggtgaggtgaggt 1499

RESULT 14
US-09-248-528-3
Sequence 3, Application US/09248528C
Patent No. 6153415
GENERAL INFORMATION:
APPLICANT: Oriol, Patrick J
APPLICANT: Padmakumar, Rugmini
APPLICANT: Kim, Sang H
TITLE OF INVENTION: Method for Producing Amide Compounds Using a Nitrile
FILE REFERENCE: MSU 4.1-401
CURRENT FILING DATE: 1999-02-10
EARLIER FILING DATE: 1998-04-29
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 1516
TYPE: DNA
ORGANISM: Bacillus pallidus
FEATURE:
NAME/KEY: rRNA
LOCATION: (1)..(1516)
PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER: Z26930/GenBank
DATABASE ENTRY DATE: 1997-05-14
US-09-248-528-3

Query Match 53.5%; Score 789.8; DB 3; Length 1516;
Best Local Similarity 75.0%; Pred. No. 3.5e-275;
Matches 1124; Conservative 0; Mismatches 327; Indels 48; Gaps 9;
9 cctcaaacacacgctggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 68
1 cctcaagacgaacgctggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 60
69 ag-----cgcgcaacggtgagtaacacgtgtgtaactcttccgc-agtct 113
61 tgcctccttaagttacagc 120
114 gggataacttcgaaagaaagcgaatcgcgatagtcgttggatcaagaatttga 173
121 gggataacttcgaaagaaagcgaatcgcgatagtcgttggatcaagaatttga 180
174 taggttaag-----attatgtcttgagatgagccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 225
181 gttgaagc 240
226 tgaagtaactgctacacaaagcgaacgtcgtgagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 285
241 tgaagtaactgctacacaaagcgaacgtcgtgagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 300

OY	286	atgtaattcgagacaacggtccatactactcttcgagagcgacgattgaagatcttctgcaat	345
Db	301	ctggagatctgagacaacggtcccaagactctcccgagagagcaatgaagatcttccgcgaca	360
OY	346	ggggagaaaccttgaagacgacgcgcgttgacaagaaagagttcttcgattgttaagtt	405
Db	361	ggagcaagaagctctgacgaggaacaacgcgcgttgagcgagaagagttcttcgaltcgtaagct	420
OY	406	catlaagcgaggaataaataagcagcaattgta-----tgatgtaacctgacct	451
Db	421	ctgtgtctcagaggaagaaacaagttgcgtctcgacaacagggcgtaacctgtgacggttaacctgacg	480
OY	452	a--aagacacgggtctaaactaagctgacgacgcgcggttaatacgtatagtgtaaaagttg	509
Db	481	agagaagccaacggtctaaactaagctgacgacgcgcggttaatacgtatagtgtaaaagttg	540
OY	510	tgcgaatcatctggcggtlaaaagtggtcgctagcgcgatcttgtaagtcagagtggtgaaactg	569
Db	541	ccggaatcatctggcggtlaaaagtcggtcgacgcgcgcgtcccttaagttctgattgtgaaactcgc	600
OY	570	gggtctcaacccggtgacctgtgacacttgaaactataaagctctggaattctgggaagagcaagtg	629
Db	601	cggtctcaacccggtgacctgtgacacttgaaacttgaaacttgaaacttgaaacttgaaacttg	660
OY	630	aattccaggtgtctgacggttgaaatgctglaagataltcttgagaaacacacacgttgagaaagcgca	689
Db	661	aattccacggtgtgacggttgaaatgctglaagataltcttgagaaacacacacgttgagaaagcg	720
OY	690	ctgtctgtgctccaacactgacgctgagagcagcagaaagcgtgtgattgaacagagatagata	749
Db	721	ctctctgtgctccaacactgacgctgagagcagcagaaagcgtgtgattgaacagagatagata	780
OY	750	cccggtgtaactcaagccctcaaaagctgtgtctacaacagttgttgtaggggtttt--aaacctag	807
Db	781	cccggtgtaactcaagccctcaaaagctgtgtctacaacagttgttgtaggggtttt--aaacctag	840
OY	808	taacgaaaccttaacgagatttaagtgtaagcgcgccttgaggagctatgtgtcgcagaaagtgaaactca	867
Db	841	tgtcgtcagcaaaacgcatctaaagcactccgcgtcgggtgagtgatgcgcgcaagctgtgaaactca	900
OY	868	aaggaatattgacgggtgtctcgcacaaagcgtgtgtagacatgtgtgttatttcgataagatacccc	927
Db	901	aaggaatattgacgggtgtctcgcacaaagcgtgtgtagacatgtgtgttatttcgataagacacgcg	960
OY	928	aaaaaacctcaacctgtggtcttgacatctgagactgtaactc----atgtagagataatatagacct	982
Db	961	aaggaacttcaacgagttcttgacatctccctcgacaacaccttagagataagagtcgttcccttc	1020
OY	993	tcgggtcagatattcaacaggtgtctgtgacatgtgtgtcgtcgtcagctcgtgtcgtgagatagttgggt	1042
Db	1021	ggggagacaagggtgtaaaagttgtgtcagatgtgtgtcgtcgtcagctcgtgtcgtgagatagttgggt	1080
OY	1043	aagctccgcgcaaacgacgcgaacacctctgctgatatgtccta--ccttaagttgttgacatgtgta	1101
Db	1081	aagctccgcgcaaacgacgcgaacacctctgacacttagttgtgcagacatcaagtgtggcactctaa	1140
OY	1102	cgaaaactgtcggtgtgacaacacggagaggaagtcgggtgacgttcaactccctcatgacctt	1161
Db	1141	gggttattgtctcggtgtgacaagagctcgagaggaaggttggtggtgacgttcaactccctcatgacctt	1200
OY	1162	atgtcgaagggccaacacacgctgtctacatattgccaatacaagaggtgtgcgcaacatcgcaaaag	1221
Db	1201	atgtcgcctgggtgtacacacgctgtctacatattggtgtgtcacaagagtcagcgaacacggtgaggt	1260
OY	1222	ggagactaactctaaaggtcgttcccaagttcggatgtgggtgtgcgaactcgaccccatgta	1281
Db	1261	ggagcgaatccccaanaaacacactccagtctgagattgtcagctgtgcacactcgtcgtcatgta	1320
OY	1282	agtcgagatcgtctagatattcgcgatcagatcagatgacgcggttgatatacgttcccggaaccttg	1341
Db	1321	agccggaatcgtctagatattcgcgatcagatgacgacgcggttgatatacgttcccggaaccttg	1380

```

OY 1342 tacacacgcccgtacacccacctgaagtgaggagacccgaagtggcttctgttaacgt 1401
Db 1381 tacacacgcccgtacacccacctgaagtgaggagacccgaagtggcttctgttaaccc 1440
OY 1402 a-agagacagactactaagtcgaacctcgtlaaagggggtgaagtcgttaacaagttacc 1459
Db 1441 acggagacagactactaagtcgaacctcgtlaaagggggtgaagtcgttaacaagttacc 1499

RESULT 15
US-09-549-108-3
: Sequence 3, Application US/09549108
: Patent No. 6214603
: GENERAL INFORMATION:
: APPLICANT: Oriol, Patrick J
: APPLICANT: Padmakumar, Rugmini
: APPLICANT: Kim, Sang H
: TITLE OF INVENTION: Method for Producing Amide Compounds Using a Nitrile
: TITLE OF INVENTION: Hydratase from a Thermophilic Bacillus
: FILE REFERENCE: MSU 4.1-486
: CURRENT APPLICATION NUMBER: US/09/549,108
: CURRENT FILING DATE: 2000-04-13
: PRIOR APPLICATION NUMBER: 60/083,485
: PRIOR FILING DATE: 1998-04-29
: PRIOR APPLICATION NUMBER: 09/248,528
: PRIOR FILING DATE: 1999-02-10
: NUMBER OF SEQ ID NOS: 18
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 3
: LENGTH: 1516
: TYPE: DNA
: ORGANISM: Bacillus pallidus
: FEATURE:
: NAME/KEY: rRNA
: LOCATION: (1)..(1516)
: PUBLICATION INFORMATION:
: DATABASE ACCESSION NUMBER: Z26930/GenBank
: DATABASE ENTRY DATE: 1997-05-14
US-09-549-108-3

Query Match 53.5%; Score 789.8; DB 4; Length 1516;
Best Local Similarity 75.0%; Pred. No. 3.5e-275;
Matches 1124; Conservative 0; Mismatches 327; Indels 48; Gaps

OY 9 ctcaagaactaacgcttgcgcgcgtcttaacaatcgaaagtcgagcggttagacaatcct 68
Db 1 ctcaagacgaacgcctgcgcgcgtcttaacaatcgaaagtcgagcggttagacaatcct 60
OY 69 ag-----cggcgaacgggtgagtaaacagcggtgaatcttcctcgcg-agtct 113
Db 61 tgcctccttaagttaaacgcgcgcgcgtggtgagtaaacagcggtgaatcttcctcgcgaact 120
OY 114 gggataactcttcgaagaagaaagctaatatccgcgatagtcctgttgatcacaaatttga 173
Db 121 gggataactcttcgaagaagaaagctaatatccgcgatagtcctgttgatcacaaatttgcg 180
OY 174 taagctaaag-----attatctgttggaatgaagcccgccgcgcgatatactagtgtg 225
Db 181 gttgaagaagcgcgtctttagctgtctacatcaggaatggtggccgcgcgcgtatgaactagtgtg 240
OY 226 tgaagtaatggtgtcccaagaagcgacgatcgcgttagccgcgcgtctagaaggtgtccgcgcaca 285
Db 241 tgaagtaatggtgtcccaagaagcgacgatcgcgttagccgcgcgtctagaaggtgtccgcgcaca 300
OY 286 atggaactgagacacggttcatactactcctacggaagcgacgtaagaaattctgttcaat 345
Db 301 ctggtgactgagacacggtcccaagactcctactcagggagcgacgtaagaaattctgttcaat 360
OY 346 ggggggaaacccctgaagcagcgacgcgcgcgtgtgaacgaagaaggtcttcgtatgttaagt 405
Db 361 ggaacgaaggtctggaagcagcgacgcgcgcgtgtgtgaacgaagaaggtcttcgtatgttaagt 420

```


THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 01:10:40 : Search time 4441.54 Seconds
(without alignments)
3573.426 Million cell updates/sec

Title: US-09-380-826A-1
Perfect score: 1477
Sequence: 1 gatcagctcgaactaac.....ccgtaacgctcctgcag 1477

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 537289281 residues
Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estfun:*
2: em_esthm:*
3: em_estln:*
4: em_estom:*
5: em_estpl:*
6: em_estda:*
7: em_estro:*
8: em_estov:*
9: em_hic:*
10: gb_estl:*
11: gb_estc2:*
12: gb_hic:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_hyr:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rtd:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Query Length	DB ID	Description
C 1	436.8	29.6	800	13	AO960400 LERFD51TR
C 2	424.8	28.8	789	12	BH145112 TDGEP51TH
C 3	424	28.7	716	13	AO957363 LERAP36TR
C 4	411.4	27.9	868	11	BF620647 HYSMC002
C 5	411	27.8	730	13	BH144909 TDGEQ40TH
C 6	407.4	27.6	720	13	CNS00WCV
C 7	392.8	26.6	755	13	BH016730
C 8	389	26.3	746	13	AO957362 LERAP36TR
C 9	383.6	26.0	741	13	AO957181 LERAP33TR
C 10	378.8	25.6	757	13	AO957182 LERAP34TR
C 11	376.6	25.5	676	13	B73824
C 12	364.8	24.7	655	13	BH016557 TDGCN35TH

C 13	360.4	24.4	693	13	AO956298 LERAJ19TR
C 14	359.6	24.3	643	13	BH144673
C 15	359.6	24.3	676	13	AO969911 LERJ28TR
C 16	359.4	24.3	643	13	BH007227
C 17	356.6	24.1	659	13	AO969996 LERUR90TR
C 18	354.4	24.0	716	13	AL153840 Anopheles
C 19	352.4	23.9	677	10	AW221851 EST298662
C 20	351.4	23.8	624	13	BH145348 TDGEV44TH
C 21	350.6	23.7	616	13	AG019017 Homo sapi
C 22	346.8	23.5	675	13	AO956297
C 23	343	23.2	603	10	AV531914
C 24	342.2	23.2	655	11	BI207639 EST525679
C 25	340.4	23.0	768	13	AO957272 LERAO84TR
C 26	340	23.0	637	13	AO962145 LERGD59TR
C 27	339.6	23.0	687	13	AO955848 LERAG49TR
C 28	339	23.0	632	13	AO967817 LERIV15TR
C 29	337.4	22.8	618	13	AO969912 LERUR28TR
C 30	336.6	22.8	601	13	B27166
C 31	335.6	22.7	775	11	BG645215 EST506834
C 32	334.2	22.6	660	10	AM600903 3B5 CDNA
C 33	333.6	22.6	592	13	AG019018 Homo sapi
C 34	332.4	22.5	599	13	AO968703 LERJ531TR
C 35	332	22.5	648	13	AO957342 LERAP25TR
C 36	331.6	22.5	561	13	AG019935 Homo sapi
C 37	331.6	22.5	616	10	BE356988 DGI_146_B
C 38	331.6	22.5	673	13	AF075762 AF075762
C 39	330.8	22.4	604	10	BE204135 EST396811
C 40	327.6	22.2	662	10	AW736167 EST332153
C 41	327.2	22.2	639	13	AO956183 LERAI55TR
C 42	325.6	22.0	559	11	BF845685 780B5 CDN
C 43	324.4	22.0	610	13	AO961463 LERFL40TR
C 44	323.8	21.9	739	13	AL146887 Anopheles
C 45	319.8	21.7	553	5	BE362383 DGI_86_HO

ALIGNMENTS

RESULT 1
AO960400/c 800 bp DNA GSS 28-JAN-2000
LOCUS LERFD51TR LERA Arabidopsis thaliana genomic clone LERFD51, DNA
DEFINITION
sequence.
ACCESSION
AO960400
VERSION
AO960400.1 GI:6788101
KEYWORDS
GSS.
SOURCE
tiale cress.
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidops.
REFERENCE
1 (bases 1 to 800)
Buell,C.R., Lin,X., Pai,G., Barnstead,M., Bowman,C., Uterbach,T.,
Feldblum,T., Liang,F., Creasy,T. and Fraser,C.M.
Genomic survey sequencing of Landsberg erecta ecotype of
Arabidopsis thaliana and identification of sequence-based
polymorphisms
Unpublished (2000)
CONTACT
Contact: Xiaoying Lin
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: atc@igrr.org
FOR ADDITIONAL INFORMATION, see <http://www.tigr.org/tdb/at/at.html>
Similar to A. thaliana chloroplast sequence (GB:AP000423)
Seq primer: TR
Class: shotgun.

FEATURES

SOURCE
1..800
/organism="Arabidopsis thaliana"
/strain="Landsberg erecta"
/db_xref="taxon:3702"

Db 483 GTACGTTGCAAGATGAACTCAAGGAATTGACGGGGGCCCCACAAAGCGGTGAGACA 424
 Qy 904 tgtgttaataatgatataaccaccaaaactcctcgtgggttgacatggtctgaac- 962
 Db 423 TGTGTTTAATTATTCAGCAAGCAAGCAACCTTACCAAGGCTTGACATGCGCGAATCT 364
 Qy 963 atcagagataatagagccttcggg--cagatcaagctcgtcatgtgttcgcagc 1020
 Db 363 CTGAAGAGAGGGGTGCTCTCGGGAACGGACACAGGTGTCATGCTGCTCAGC 304
 Qy 1021 tctgtcgtgagatgttgggttaagtcctcgaagcgcgaacctatcgtatgttgc 1080
 Db 303 TCGTGCGGTAGGTGTTGGGTTAACTCCCGACAGAGCGCAACCTCGTGTAGTTGCC 244
 Qy 1081 a-ccttaagtggcagcgtgtaagaaactcgcgttgacaaacccgagagagcggtgtg 1139
 Db 243 ATCTGTGAGTTTGAACCTTGAACAGACTGCGGTGATTAACCGGAGAAAGTGAAGATG 184
 Qy 1140 acgtcaaatccatagccttatgtccaggacacacagctgtacatagccgatatca 1199
 Db 183 ACGTCAGTCAATCATGCCCTTATGCCCCGCGACACACAGTCTACATAGCGCGGACA 124
 Qy 1200 gagggtcgcaactcgcaagagggagatctcttaaaagtcggtccagtcggtatgg 1239
 Db 123 AAGGTCGCGATCCCGGAGGGTGAAGTAAACCCCAAAACCCTGTCATGTCGATTCG 64
 Qy 1260 ggtctgcacccagcccatgaagtcggaatcgtctagtaatcgcgagatcagc 1311
 Db 63 AGGCTGCAATCGCCTCGATGAGCCGGAATCGCTAGTAATCGCGGATGAC 12

RESULT 3
 LOCUS AO957363 716 bp DNA GSS 28-JAN-2000
 DEFINITION LERAP36TR LERA Arabidopsis thaliana genomic clone LERAP36, DNA sequence.
 ACCESSION AO957363
 VERSION AO957363.1 GI:6785064
 KEYWORDS GSS.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 716)
 Buell, C.R., Liu, X., Pai, G., Barnstead, M., Bowman, C., Uterbach, T., Feldblyum, T., Liang, F., Creasy, T., and Fraser, C.M.
 Genomic survey sequencing of Landsberg erecta ecotype of Arabidopsis thaliana and identification of sequence-based polymorphisms
 Unpublished (2000)
 JOURNAL Contact: Xiaoying Lin
 COMMENT The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: atetlgr.org
 For additional information, see <http://www.tigr.org/cdb/at.html>
 Seq primer: TR
 Class: Shotgun.
 FEATURES
 source Location/Qualifiers
 1..716
 /organism="Arabidopsis thaliana"
 /strain="Landsberg erecta"
 /db_xref="taxon:3702"
 /clone="LERAP36"
 /clone_1id="LERA"
 /note="Organ: Leaf; Vector: PHOS1; Total genomic DNA was sheared to 0.9-1 Kbp before ligation."
 BASE COUNT 144 a 213 c 154 g 205 t

Query Match 28.7%: Score 424; DB 13; Length 716;
 Best Local Similarity 77.6%; Pred. No. 4,36-106;
 Matches 545; Conservative 0; Mismatches 140; Indels 17; Gaps 2;

Qy 275 gtccgcccacaaatgaactgaaacggttcatactcctacggaagcagatgaagaa 334
 Db 716 GACCAAGCCACATGAGAACTAGACACGGTCCAGACTCTTACGGAGGACAGTGGGAA 657
 Qy 335 tctgtctaaatgggggaaacccctgaagcagcagccgcgctggaacgaagagcttcgg 394
 Db 656 TTTTGGACAATGGGGCAAGGCTTATCCAGCATTCCTCCGGAGTGAAGAGGCTTCGG 597
 Qy 395 attgaaagttcatataggcaggaanaataagcaagcaatgatgatgttaccctca- 452
 Db 596 GTTGAAGCTCTTTCCGAAGGAAAGAACTTACTTTCTATAAGAGTGAAGCTGACG 537
 Qy 453 -----aagcacggtctaactcagtcgcagcagccgcggtatacgtatgtg 499
 Db 536 GTACCTGATTAAGAAAGCACCGGCTTAAGTGCACAGCCCGGTAAATACGTAGGGTG 477
 Qy 500 caagcgtgttcggaatcatatgggcgttaaggtgctgtagcgatgtgtaagtcagtg 559
 Db 476 CGAGGTTAATCGGAATTAATCTGGGGGTAAAGCGTCCGAGCGGTTGGCAAGTCAGATG 417
 Qy 560 tgaacactgcggtcgaacccgtggtcgtcaactgaaactaactcaagctcgtgagttggag 619
 Db 416 TGAATATCCCAAGCTCAACTTGGGAACTGCTTTCAAACTCCGACACTAGATATGTGAG 357
 Qy 620 aggcagtggaattccaggttagcgtggaatgctgtagatatactggaagaaacccagtg 679
 Db 356 AGGGGGGTAGAAATTCACAGTGTAGCAGTGAAGATGCGGTAGATGTGAGATACCAATG 297
 Qy 680 gcgaagcagactgtgtgtcctaaactcagctgagcagaagaagcgtgtgtagtaacg 739
 Db 236 CGGAAGCGAGCCCTCGGATTAATTTGACGCTCTATGACCAAGACGGGGAGCAAAACA 237
 Qy 740 ggaatagatacccccgttaataccaagccctaaacgtgtgtcacaagttgttgggtttta 799
 Db 236 GGATTAGATACCTCGGTGATGTCACGCCCTTAAGCATGTCTAGTGTGGTAGTA 177
 Qy 800 a--ccctcagtaacgaacctaacgattagtagacgcgttcgtggactatcctcgcaaga 857
 Db 176 AATCATATAGTAGACACACACTTAACGGTGAAGTAGACCCCTCGGGAGATACGCTGCAAGA 117
 Qy 858 gtgaacactcaagaagatagcaggggttcgcacaaagcgtgtgagatgtgttatttcg 917
 Db 116 TTAATACTCAAGGAATGACGGGGGCCGCAAGACGGGTGATTATGTGATTAATTCG 57
 Qy 918 atgatacccaaaaacacccactcgtggttgacatgacatga 959
 Db 56 ATGCACGCGAANAACCTTACCTGCGCTTGACATGCCACTAA 15

RESULT 4
 LOCUS BF620647 868 bp mRNA EST 22-FEB-2001
 DEFINITION HVSMEC0020H16f Hordeum vulgare seedling shoot EST library
 HVCDNA0003 (Ptilated and unstressed) Hordeum vulgare cDNA clone
 HVSMEC0020H16f, mRNA sequence.
 ACCESSION BF620647
 VERSION BF620647.2 GI:13109622
 KEYWORDS EST.
 SOURCE barley.
 ORGANISM Hordeum vulgare
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
 ; Triticeae; Hordeum.
 1 (bases 1 to 868)
 Wing, R., Close, T.J., Kleinhofs, A., Wise, R., Begum, D., Frisch, D., Yu, Y., Anderson, H., Dale, J., Henry, D., Kennolly, S., Palmer, M., Rambo, T., Saski, C., Schwartzbeck, V., Simmons, J., Choi, D.W., Main, D. and Wood, T.
 Development of a genetically and physically anchored EST resource
 TITLE

OY	659	atactctggagagaaacccagtgctgcgaagagcacttgctgcctctaaacactgacgtctgaagca	718
Db	670	AGATGGAAAGAACACCAACGCGGAAAGCACTCTGCTGGCCGACACTACACTGTGAGAA	611
OY	719	cgaagcgtctggatgaataacggatctagaatacccgctataccacgcccctaaacgtctgc	778
Db	610	CGAAAGCTAAGGGAGACGAATGGGATTAGATTACCCAGTAGTGCTCTAGCCCTTAACAGATGGA	551
OY	779	taccagttgctgggggctttaaacc--tcagtaacgaacactaaggatlaagtagacgcg	836
Db	550	TACTAGGGCGCTGTGGTTCGACACCGGTGACGTCTACTTAACCGGTTAAGTATCCCCC	491
OY	837	ctgggggactatgctctgcgaagtggaacactcaagaagatagacgggggtctgcacaaagcg	896
Db	490	CTGGGGAGTACCTTGCAAGATGAAGTCAAGGATTGACGGGGGCCCGGACAAGCGG	431
OY	897	tggagcatgtggttctaactcgaatgaatacccaaaaacactcaccttggtctgacatgatac	956
Db	430	TGGACCATGTGGTTTAATTCGATGCAGAAACGAAACCTTACGAGGGCTTGACATGCCG	371
OY	957	tgaatc-atgtagaatatatagaaccttcgggcagat--tcacagttgctcatgtgtt	1013
Db	370	GAATCCTCTTGAAAGAGAGGGGTGCTTCGGGAACGCGGACACAGTGGTGCATGGCTTT	311
OY	1014	cgctacgctcgtctcgtgagatgttggttaagtccgcgaagcgcaacccctatcgta	1073
Db	310	CCTCAGCTCGTCCCTTAAGGTCTTGCGTTAATCCCGGACACGACCCCAACCTCGTGT	251
OY	1074	tgttgcta-ccctaaagttgggacactgtaacgaacactgcgcgttgacaacacggagaaagcg	1132
Db	250	AGTTTCCATCGTTGAGTTTGGAAACCTTGAAACAGACTGCGGTTGATTAAGCCGAGGAAGGT	191
OY	1133	ggggatgacgttaaatctctcatagccttatgttccaagggccacacacgtgtctacaatggc	1192
Db	190	GAGGATGACGCTAACATCATCATGCCCTTATCCCTGCGGCGACACACGTCATCAATGGC	131
OY	1193	cgatacgaagggctgcgaacactcgcaagagagagactcaatctctaaagtcgtgcccaattc	1252
Db	130	CGGGCAAAAGGTCCTCGATCCCGGCAAGGCTGAGCTAACCCCAAAAACCTGCTCTCATGTT	71
OY	1253	ggattggggtctctgaactcgaccccatgaagtcgggaatcgtctagtaactcgcgcatcagc	1311
Db	70	GGATTGCAAGGCTGCAACTGCGCTGCATGATGAAGCCGGAATGCTAGTAATGCGCGGTGAGC	12
RESULT	6		
CNS00MCV	720 bp	DNA	GSS
LOCUS	28-JUN-1999		
DEFINITION	Arabidopsis thaliana genome survey sequence SP6 end of BAC T11H14		
KEYWORDS	of T11H14 library from strain Columbia of Arabidopsis thaliana,		
ACCESSION	genomic survey sequence.		
VERSION	AL093421		
KEYWORDS	AL093421.1 GI:5294575		
SOURCE	GSS.		
ORGANISM	thale cress.		
REFERENCE	Arabidopsis thaliana		
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
JOURNAL	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;		
TITLE	Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.		
FEATURES	1 (bases 1..to 720)		
FEATURES	Salanoubat,M., Choiane,N., Artiguenave,F., Brottier,P., Wincker,P.,		
FEATURES	Samson,D., Saurin,W., Weissenbach,J. and Quetier,F.		
FEATURES	Unpublished		
FEATURES	2 (bases 1 to 720)		
FEATURES	Genoscope.		
FEATURES	Direct Submission		
FEATURES	Submitted (25-JUN-1999) Genoscope - Centre National de Sequencage :		
FEATURES	BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr		
FEATURES	- Web : www.genoscope.cns.fr		
FEATURES	Location/Qualifiers		
FEATURES	1..720		
FEATURES	/organism="Arabidopsis thaliana"		

Query Match	27.6%	Score 407.4	DB 13	Length 720
Best Local Similarity	77.2%	Pred. No. 1.6e-101		
Matches	547	Conservative	0	Mismatches 156; Indels 6; Gaps 4
QY	609	gaggttggaagaggaatggaattccagggtgtagcgggtgaaatcgtaatatctcgag	668	
Db	1	GAGTACGGTTGGGGCGAGAGGAATTTCCGGTGGAGCGGTAAATGCGTAAAGATCGGAAA	60	
QY	669	gaacacagctgaggaagaggaacttgcttggtcgaacactgacgcgtggaagcagaaacgtg	728	
Db	61	GAAACACCAACGGCGAAGAGCACTGCTGCTGGGCCACACACTGACACTGAGAGAGCAAGACTAG	120	
QY	729	ggttagtaaacgagatagatagatcccggttaatccagcgcctaaacgltgtctaccagttgt	788	
Db	121	GGAGCGGATGGGATAGATAGATACCCCGATAGTCTTACGCCGTAAGAGATGATAGACGCG	180	
QY	789	tgggggtttaaacc--tcagttaacgaactaacgataagtagaacgcgttggggagcta	846	
Db	181	TGCGCTATCGACCCCTGACAGTCTCTAGCTTAACGGCTTAAGATCCCGCTGGGAGTA	240	
QY	847	tgctcgcaagagtgaaactcaaaagaaatgaaggggtccgcaaaacggttgagacatgt	906	
Db	241	CGTTGCCAAGATGAATCACTCAAGAGATTACCGGGGGCCCGCACACGGGTGAGCATGT	300	
QY	907	ggtttaactcgatgataccccaacaaactccaccctgggcttgatgatatgtaatcatgt	966	
Db	301	GGTTTATTTTCGATGCAAGCGAAGAACTTACACAGGCTTGACATGCCCGGAATCCTCTT	360	
QY	967	-agagatatgagccttcgggcagat--tcacagtgctgcatgltgttcgtacgtcg	1023	
Db	361	GAAAGAGAGGGGTCCTTCGGGGAACCGGACACAGGCTGTGCTGTCGTACGCTCG	420	
QY	1024	tgtctggaatatgttgggttaagtcccgcaacgagcggaacccctatcgtatgtgttgcacc	1083	
Db	421	TGCGCTAAGAGTGTGGGTGAAGTCCGCAACGAGCCGACCCCTGTGTTTAAAGTTCACAC	480	
QY	1084	-ttaagtctggacactgtaagaaactgcgcgtgtgacaacacccggaagagcgcgagatagc	1142	
Db	481	GTTGAGTTTGGAAACCTCGAACACAGACTGCGGTTATAGCCGGAGAGAGTGATGAGTACG	540	
QY	1143	tcaaatctcatatgaccttattgtccaggggcacacagcgtgctacaatlygcogatacagag	1202	
Db	541	TCAAGTCATCATGTCCTTCCTTATGCCCTGGGGACACACAGTCTCAAAATGGCGCGGACAAAG	600	
QY	1203	ggttcgcaactcgcgaagggagagctaatctcttaaaagtctgtccagttcgtgatttgggt	1262	
Db	601	GGTCGGATGTCGCCGAGGGGAGCTTAATCTCAAAACCCCTTCACTTGGATGGATGGAGC	660	
QY	1263	ctgcaactgcagcccaatgaagtgcgaatcgtagtaatcgcgatcagc	1311	
Db	661	CTGCACTGCGCCCTGCATGAAAGCCGGAATCCCTAGTATGCCCGGTCAAGC	709	

RESULT 7
 BH016730/c
 LOCUS
 DEFINITION TDGCRP2TH cTog Lycopersicon esculentum genomic clone cTog17F2, DNA
 sequence.
 ACCESSION BH016730
 VERSION BH016730.1 GI:14147768
 KEYWORDS GSS.
 SOURCE tomaato.
 ORGANISM Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

QY	204	ccggcgccgcatagctatgcttggtgagtgtaatgctcaacaaagcgacgctgctgtaacgg	263
Db	119	CGATATCTGATTACTAGTGGTGGGTAAAGGCTTACCAAGGCTTCCGACTCATGACTACTGG	178
QY	264	ccttagaggggtgtccggcccaatgtgagactgagacacgctcactactctacggagagca	323
Db	179	TCTGAGAGGAGACGACCACGACACTGGAAGTACAGACGGTCCGAGCTCCGACGGAGGCA	238
QY	324	gcagtttaagaactctgtctcaatgtgggggaaaccttgaaacagcagcgccgctgtaacgaag	383
Db	239	GCACTGGGGGAATTTTGGACAATGGGCGCCCAAGCCCTGATCCAGCCATTCCCGTACGGAAG	298
QY	384	aaggtctcggatgtgtaaatgtaatcattagcaggagaaataagcagcaatgtgatagtt	443
Db	299	AAGGCTTCCGGGTTGTAAAGCTCTTTTCGCAAGGAGGAGAAACTTACTTCTATTAAGAG	358
QY	444	acctgacctt-----aagcacccgctactactcgtgccaagcaccgggttaa	488
Db	359	TGAGCGTGCAGCGTACTTGTATAGAAACACCGGCTAACTACGTGCCACGACGCGCGGTAA	418
QY	489	tagctatgtgtgcaagcgtctgtcggaaatcattggcgctaaaggctgcgtagcgagatttg	548
Db	419	TACCTAGGCTCCGAGCCTTATCGGAATTACTGGGCTTAAGCGCTGCGCAGCGGCTTTGG	478
QY	549	taagtcaggtgtgaaacatcgcggtctcaaccgctggtcgtcacttgtaaatcaactaagctcg	608
Db	479	CAACTCAGATGTGAAATCCCGAGACTCAACTTGGGAATCTCGTTGAAATCGCCAACTA	538
QY	609	gagcttggagagagcgaagtgtgaattccaggtgtatgacgtgtgaatgtgtatctcggag	668
Db	539	GAATATGTCAAGAGGGGGGTAGAAATCCACGTGTAGCGTGAATATGCGTAAAGTAGTGAG	598
QY	669	gaacacacgtgtgcgaagcgacactgtcgtctcaaaacttgacgtctgagcgacgaagacgtg	728
Db	599	GAATACCAATGCGCAAGGACGACCCCGCTGNGATATATATGACGCTCATGCAGAAACGTG	658
QY	729	ggtgttaaaacggatctgatacccggttaatccaacgcgccttaaacgtctgtctacagttgt	788
Db	659	GGGAGCAAAACGAGTATGATACCTGGTATGCCACGCCCTTAACAGATGCTACTAAGTTGT	718
QY	789	tggggtgttttaa 800	
Db	719	TGGTGCAGATAA 730	
RESULT	9		
QY	A0957181/c	741 bp DNA	GSS 28-JAN-2000
LOCUS	LERA033TR LERA Arabidopsis thaliana genomic clone LERA033, DNA		
DEFINITION	sequence.		
ACCESSION	A0957181		
VERSION	A0957181.1	GI:6784882	
KEYWORDS	GSS.		
SOURCE	thale cress.		
ORGANISM	Arabidopsis thaliana		
REFERENCE	Eukaryotes: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta		
AUTHORS	Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots;		
TITLE	Rosidae; eustroids II: Brassicales; Brassicaceae; Arabidopsis.		
	1 (bases 1 to 741)		
	Buell,C.R., Lin,X., Pal,G., Barnstead,M., Bowman,C., Uteirbach,T.		
	Feldjbaum,T., Liang,F., Cressy,T. and Fraser,C.M.		
	Genomic survey sequencing of landsberg erecta ecotype of		
	Arabidopsis thaliana and identification of sequence-based		
	polymorphisms		
JOURNAL	Unpublished (2000)		
COMMENT	Contact: Xiaoying Lin		
	The Institute for Genomic Research		
	9712 Medical Center Dr., Rockville, MD 20850, USA		
	Tel.: 301 838 0200		
	Fax: 301 838 0208		
	Email: atetlgr.org		
	For additional information, see http://www.tlgr.org/tdb/at/at.html		
	Similar to A. thaliana chloroplast sequence (GB:AF000423)		

FEATURES	source	Seq primer: TR	Class: shotgun.	Location/Qualifiers
				1. .741
				/organism="Arabidopsis thaliana"
				/strain="Landsberg erecta"
				/db_xref="taxon:3702"
				/clone="LERO33"
				/clone_11p="LERA"
				/note="Organ: Leaf; Vector: pHOS1; Total genomic DNA was sheared to 0.9-1 Kbp before ligation."
BASE COUNT	137 a	240 c	169 g	195 t
ORIGIN				
Query Match	26.0%	Score 383.6;	DB 13;	Length 741;
Best Local Similarity	74.6%	Pred. No. 5.8e-95;		
Matches 525;	Conservative 0;	Mismatches 169;	Indels 10;	Gaps 3;
OY	190	gcttgagatcagcccgccgagcagatcagctgattgtgtagagtaatgctccacaagcgca	249	
Db	697	GCCCCAGAGAGGGCGCTCGCGGTCTGATTACTAGTGTGTAAGCAATAGCTTACCAAGCGGA	638	
OY	250	cgatctgtatgcgcgcgcttgaagaggtgtccgcgcacaaatggaacttgaagacagcttcatac	309	
Db	637	TGATCAGTAGCTGGTCCAGAGAGATGATCAGCCACACCTGAGACTGAGACAGCGCCCAAC	578	
OY	310	tccttcacggaggaagcagcttaagaatctgtctcaatctggggaaaccctgaaagcagcagc	369	
Db	577	TCTCTAGGGAGGAGGACAGTGGGAAATTTCCCAATGGGCGAACCCTGACGGAACATG	518	
OY	370	ccgcgtgaacgaagaaggctctcgcgattgtlaaagtctlaagcctlaagcagaaataaagcagc	429	
Db	517	CCGCGTGGAGGTAGAAGGCCCTACGCGGTCTGTAACCTTTTCCAG--AGAAGAAGCAAT	460	
OY	430	aatgtgatatgtatccttcgctctaaagacccggtctactacgtgtccagcagccgcggtat	489	
Db	459	GACGGTATCTGGGGA-----ATAAGCATTCGGCTTAACCTGTGGCCAGAGCCGCGGTAT	406	
OY	490	acgtatgtgtcaagcgttgtctcgaatcatctgggcgtlaaaggtgtcgtlagcgcgaattgt	549	
Db	405	ACAGAGGATGCCAAGCGGTATCCCGGAATGTTGGCGGTAAAGCGTGTGATAGTGCTTTT	346	
OY	550	aagtcaggtgtgaaactcgcggtcctcaaccggtgtgcctgtcacttgaactataagttcgt	609	
Db	345	AAGTCCCGCGTCAATCCAGGGCTCAACCCCTGACAGCGCGGTGGAACCTACCAAGCTTG	286	
OY	610	agttgggaggaagcagcgaatctccaggtgttagaggtgtgaaatgtgtatgtatcgtgaag	669	
Db	285	AGTACGGTAGGGGCGAGAGGGAATTTCCGCTGGAGCGGTGAATTCGCTGATGAGATCGGAAG	226	
OY	670	aacaccagctgcgaagcgacttctgtgcctcaaaactgaagcgtgaagcagaagaagcgtgg	729	
Db	225	AACACCAACGCGGAAGACACTCTCTGGGCCACACTGACACTGTGAGAGCAAGAACCTTAGG	166	
OY	730	gtatgaaacgggattagatcccggtgaatcacagcccttaaacggtgtgtctacagattt	789	
Db	165	GGAGGGAATGGGATTAGATACCCAGTAGTCTTACCCGTAACGATGGATCTAGGCGCT	106	
OY	790	gggggttttataccc--taaglaacgaactcaacgattlaagtagacgcgcttgggagact	847	
Db	105	GTGCGTATTCGACCCGTCGAGTGCTGTAGCTTAACGCTTAAGTATCCCGCTGGGAGTAC	46	
OY	848	gctgcgaagagtgaactccaagaagattgaagcgggtgtccgcaca	891	
Db	45	GTTTCGCAAGAAATCAAGGAATTGACGCGGGCCCTTACA	2	
RESULT 10				
LOCUS	AO957182	757 bp	DNA	GSS
DEFINITION	LERO334TR LERA Arabidopsis thaliana genomic clone LERO334, DNA			28-JAN-2000
				sequence.

ACCESSION	A0957182
VERSION	A0957182.1 GI:6784883
KEYWORDS	GSS.
SOURCE	thalie cress. thaliana
ORGANISM	Arabidopsis thaliana
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Eudicotyledons; Magnoliophyta (eudicotyledons); core eudicots; Rosidae; eurosoids II; Brassicales; Brassicaceae; Arabidopsi. 1 (bases 1 to 757)
AUTHORS	Buell,C.R., Lin,X., Pai,G., Barnstead,M., Bowman,C., Utterbach,T., Reidblyum,T., Liang,F., Cressy,T., and Fraser,C.M. Genomic survey sequencing of Landsberg erecta ecotype of Arabidopsis thaliana and identification of sequence based polymorphisms
TITLE	Unpublished (2000) Contact: Xiaoying Lin The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 838 0200 Fax: 301 838 0208 Email: atc@tigr.org For additional information, see http://www.tigr.org/db/at.html Similar to A. thaliana chloroplast sequence (GB:AP000423) Seq primer: TR Class: shotgun
JOURNAL	Location/Qualifiers
COMMENT	1..757 /organism="Arabidopsis thaliana" /strain="Landsberg erecta" /db_xref="taxon:3702" /clone="LERA034" /clone.lib="LERA" /note="Organ: Leaf; Vector: pHOS1; Total genomic DNA was sheared to 0.9-1 kbp before ligation." BASE COUNT 137 a 247 c 170 g 203 t
ORIGIN	
Query Match	25.6%; Score 378.8; DB 13; Length 757;
Best Local Similarity	74.6%; Pred.No.1.2e-93;
Matches 519; Conservative	0; Mismatches 167; Indels 10; Gaps 3.
OY	190 gccctgagatgtaccccgccgcgatctagtcttgtagtaattgctcaccaaggcga 249
DB	689 GCCCGAGGAGGGCTGTGCGTCTATTAGTAGTGTTGGTGAAGCCAAATAGCTTCAAGGCGA 630
OY	250 cgatcgtagtcggcgcttgtaggggtgtccggccacaatgtgaactgagaacggtccatac 309
DB	629 TGATCAGTAGTCGTGCTCCGAGCATATCAGCCACCCTGGAGTGAACAAGCCGCCACGAC 570
OY	310 tctctagggagagagcagtttaagaatcttgtctaactcggggaaacctgaagcagcgacg 369
DB	569 TCTTACGGAGGCACGACTGTGGGAATTTTCCGAATGGGGGAAGCCTGACGAGCAATG 510
OY	370 ccgctgtgaacgaagaagttctcgattgttaaagttcataagcgagaaaaataagcagc 429
DB	509 CCGCGTGAGAGTGAAGAGCCCTACGGGTCCTGAACCTCTTTTCCAG--AGAAGAACAAT 452
OY	430 aatgtgtagtgtlactctctaagaaccggcttaactaactgtagcagcaacgctgtaat 489
DB	451 GAGCGATTCGGGGA-----ATTAAGCATGGGCTACTCTGTGCAACACCGCGGTAAT 398
OY	490 acgtatgtgtcaaacgctgtgtcggnaatcatctggcgtaaaaggtgcgtaagcgagattgt 549
DB	397 ACAGAGGATGCAACGCTTATCCGGAATGATTTGGCGCTGAAGACGCTGTAGTGAGCTTTTT 338
OY	550 aagtcaggtgtgaactactcggtcgtcaaccgcgtgctcgtcaactggaactcgaagaagcttg 609
DB	337 AAGTCCCGCCGCAATCCACAGGCTCAACCTTGACAGCGCGTGGAAATCAACAGCTTG 278
OY	610 agtttggaaggggaagtgtgaattccaaggtttaagcgttgtaaatcgttatgatattggag 669
DB	277 AGTACGGTAGGGGAGAGGGAATTTCCGGTGTGACACGGTGAATTCGTAAATCGCTAAGATCGGAAG 218

QY	670	aaacacgctggcgaagcgacgtctgcgtgcctcaaaactgtgacgtgaagcaggaagctgg	729
Db	217	AACACGACACGGCGAAAGCCTCTGCTGGCCGACACTGACACTGACAGACGAAAGCTAG	158
QY	730	gtatgaacggatlatagatacccggttaatcaccgaccttaaacgctgtctaccagttgt	789
Db	157	GGAGCGAATGGGATTAATACATACCCAGTAGCTCTAGCCGTTAAAGATGATTAAGCGCT	98
QY	790	ggggggtttaacc--tcagtaacgaacctaaaggaatgaatgtagaccgctgggacct	847
Db	97	GTGGGTATCGACCCGTCAGTGTGATACGCGTTAATGATATCCCGCTGGGAGTAC	38
QY	848	gctcgaagatgaactcaagaagatgtacggggg	883
Db	37	GTTCCGAAGATGAACACTCAAGGATTTACGGGG	2
RESULT	11		
B73824		B73824	676 bp DNA
LOCUS		T27K4R	TAMU Arabidopsis thaliana genomic clone T27K4, DNA
DEFINITION			sequence.
ACCESSION		B73824	
VERSION		B73824.1	GI:2712975
KEYWORDS			GSS.
SOURCE			thale cress.
ORGANISM			Arabidopsis thaliana
			Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
			Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
			Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE			1 (bases 1 to 676)
AUTHORS			Rounsley,S.D., Field,C.E., Bass,S., Linher,K., Linher,K., Golden,K.,
			Berry,K., Granger,D., Suh,E., Wible,C., Adams,M.D. and Venter
			,J.C.
TITLE			A BAC End Sequence Database for Identifying Minimal Overlaps in
			Arabidopsis Genomic Sequencing. Update 3
JOURNAL			Unpublished (1997)
COMMENT			Other_GSSs: T27K4TF
			Contact: Steve Rounsley
			Department of Eukaryotic Genomics
			The Institute for Genomic Research
			9712 Medical Center Dr., Rockville, MD 20850, USA
			Tel: 301 838 0200
			Fax: 301 838 0208
			Email: rounsley@tigr.org
			Seq primer: M13 Reverse
			Class: BAC ends
FEATURES			High quality sequence stop: 676.
source			location/qualifiers
			1..676
			/organism="Arabidopsis thaliana"
			/strain="Columbia"
			/db_xref="taxon:3702"
			/clone="T27K4"
			/clone_1ib="TAMU"
			/sex="thermaphrodite"
			/note="Vector: BelobACII; site_1: HindIII; site_2: HindIII
			; Produced by Rod Wing
			128 t 1 others
BASE COUNT			170 a 164 c 213 g 128 t
ORIGIN			
Query Match			25.5%; Score 376.6; DB 13; Length 676;
Best Local Similarity			76.2%; Pied No. 4.8e-93;
Matches 515; Conservative			0; Mismatches 155; Indels 6; Gaps 4;
QY	610	aatgttggaagaagatcggaatccacgtgtagcgtgtgaacatgcatgatatctgag	669
Db	1	AGTACGCTAGGCGGAGAGGAATTTCCGCTGGAGAGCGTGAATAGCTAGAGATCGGAAG	60
QY	670	aaacacgctggcgaagcgacgtctgcgtccaaactgacgcctgagcgacgaagcgctgg	729

Db 61 AACACCAAGCGGAGACACTCTGCTGGCCGACACTGACACTGAGAGACGAAGCTAG 120
Qy 730 gtaagtaacggagatagatagaccgggtatccacagccctaaccgtgtctaccagttgt 789
Db 121 GGAGCGAATGGGATTAGTACCCTCCAGTAGTCTTACCCTTAACGATGATCTAGGCGCT 180
Qy 790 ggggggtttaaacc--taagtaacgaacctaaccgattaaagtaagaccgccttgggagact 847
Db 181 GTGCCTATCGAACCCGTGACAGTCTGTACTTAACCGCTTAAGATCCCGCTGGGAGTAC 240
Qy 848 gctgcgaagatgtaaacctcaagaagattgacgggggttccgcacaagcggttggagcatgtg 907
Db 241 GTTCGCAAGAAATGAACCAAGCAATGACGGGGCCCGCAACAGCGGTGAGCATGTG 300
Qy 908 gtttaattcgatgatacccaaaaacctcaacttgggttgcatagtatcgaatc-atgtc 966
Db 301 GTTTAATTGATGCAACGAGAACCTTACAGGGCTTGCATGCGCGGAATCCTCTTG 360
Qy 967 agagataatgagcccttcggg--cagattcaacaggtgctgatagtgttcgtcaagctgt 1024
Db 361 AAAGAGAGGGGGCTTGGGAGACGGACACAGGTGGTGCATGCTGTCCACACTGCT 420
Qy 1025 gtcgtgagatgttgggttaagtccgcgaacgagcgcaaccctatcgtatgttctacc- 1083
Db 421 GCCGTAAAGCTGTGGGTTAAGTCCCGCAACGACGCCAACCCCTCGTTAGTCCACCG 480
Qy 1084 ttaagtgggacacgtgtcgaacatcgccggtgacaaacggagaggaagcgggagatgact 1143
Db 481 TTGAGATTGGAAACCTTGAAACAGACTGCCGGTCAATAGCCGGAGAGGTGAGACATGAGCT 540
Qy 1144 caaatccatagcccttatttccagggccacacagctgtcacaatggccgataagag 1203
Db 541 CAAGTCATCATGCCCTTATATGCTCTGNGCAGACACAGTGTCAATATGGCCGGGACAAAG 600
Qy 1204 gtcgcgaactcgcaagagaggaactaactctaaagtcggtccagttcgatgtgggtc 1263
Db 601 GTCCGGAATCCCGCGAGGGGTGAGCTTAACCTCAAAAACCCGCTCTGATTCGATTCAGAGC 660
Qy 1264 tgcgaactgaccccat 1279
Db 661 TGCACCTCGCTGCAT 676

RESULT 12
BH016557/c 655 bp DNA GSS 18-MAY-2001
LOCUS T06CN35TH cT0G Lycopersicon esculentum genomic clone cT0G17E22, DNA
DEFINITION sequence.
ACCESSION BH016557
VERSION BH016557.1 GI:14147595
KEYWORDS GSS.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asterales; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 655)
AUTHORS van der Hoeven,R., Sun,H., Cho,J., Utterback,T., Romning,C. and Tanksley,S.
TITLE Tomato Demethylated Genomic DNA sequences
JOURNAL Unpublished (2001)
COMMENT Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.clemson.edu/orders/index.html>
tomato demethylated genomic DNA
Seq primer: M13F-R
Class: shotgun.
FEATURES
Source 1..655 Location/Qualifiers
/organism="Lycopersicon esculentum"

/cultivar="E6203"
/db_xref="taxon:4081"
/clone="cT0G17E22"
/clone_lib="cT0G"
/tissue_type="young leaves"
/dev_stage="12-14 weeks post harvest"
/lab_host="E.coli JM109"
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; This library was made from short EcoRI digested
fragments of the genome of Lycopersicon esculentum ligated
into pBS (SK-). The fragments were cloned into the
methylation restrictive E.coli strain JM109 with the
purpose of enriching the library for non-methylated DNA
fragments. This procedure may enrich the pool of cloned
fragments in JM109 cells for sequences representing
expressed genes. Average insert size 1.27 kb."
BASE COUNT 125 a 203 c 168 g 159 t
ORIGIN
Query Match 24.7% Score 364.8; DB 13; Length 655;
Best Local Similarity 77.0%; Pred. No. 8.4e-90;
Matches 496; Conservative 0; Mismatches 142; Indels 6; Gaps 4;
Qy 674 ccagtgaggaagcgactgtgtgtcctcaaacctgacgtctgagcgcaagaagcgtgtgtag 733
Db 655 CCAAGCGCGAAMGACACTCTGCTGGGCCGACACTGACACTGACGAAAGAACTAGGGAG 596
Qy 734 taacggagatagatacccggttaatccagcgccctaaacgtgtctaccagttgttggg 793
Db 595 CGAATGGGATTGATTAACCCCACTAGTCTTACCGCTTAACGATGATGATGAGCGCTGTGC 536
Qy 794 gtttaaccc--tcagtaacgaacttaacgattagtagagacgcgtgtggactatgtc 851
Db 535 GTATGACCCGTCGTCGTAGCTAACGCGTAAAGTATCCCGCTGGGAGTACGTTTC 476
Qy 852 gcaagagtgaacctaagaagtaacgaggttcgcacaacggttggagcatgtgtt 911
Db 475 GCAGAATGAAMATCAAAAGAAATTGACGGGGCCCGCAACAGGTGAGCATGTGGTTT 416
Qy 912 aattcgtatgataccccaaaaccctacactggtgtgacatgtatc-atgtagag 970
Db 415 AATTGATGCAAGAGCAAGAACCTTACACAGGCTTGACATGCCGGAATCCTTGAAG 356
Qy 971 atatatgagcccttggg--cagattcaacaggtgtgtgtgtgtgtgtgtgtgtgtgtgt 1028
Db 355 AGAGGGGTGCTTCGGGAACCGGACACAGGTGTGATGCTGTGATGCTGTGCTGTCG 296
Qy 1029 tgaatgttgggttaagtcctccgaacgagcgcaaccctatcgtatgttcta-ccctaa 1087
Db 295 TAAAGCTGTGGGTAAAGTCCCGCAACGAGCGCAACCCCTGCTTAAATTGCCATGCTTGA 236
Qy 1088 gttggcacctgtgtacgaaactgtcggtgacaaaaccggaaggaagcggtgataagtcata 1147
Db 235 GTTTGGAACCCCTGAACAGACTGCGCGTATTAAGCCGGAGGAAGGTGAGATGACCTCAAG 176
Qy 1148 tctcatggccttatttgcagagggccacacagctgtctcaatgagcgagataagaggtgtc 1207
Db 175 TCATCATGCCCCCTTAATGCCCCGTGGGACACAGCTGTCAATATGGCCGGGACAAAGGTTG 116
Qy 1208 ccaactcgcaagaggaagctactctctaaagtcggtccagttcgtatgttgggtgttca 1267
Db 115 CGATCCCGGAGAGGTGAGGTAAACCCCAAAAACCCGTCCTCAATTGGATGAGCTGCA 56
Qy 1268 actgagcccatgaaagtggaaatcgctatgtaacgcgcatcagc 1311
Db 55 ACTCGCTGCATGAAGACCGGAATTCCTAGTAAATCCCGGTAGC 12

RESULT 13
AQ956298 693 bp DNA GSS 28-JAN-2000
LOCUS LERAJ19TR LERA Arabidopsis thaliana genomic clone LERAJ19, DNA
DEFINITION

Oy	752	ccggatcaacacgcgcctcaaaacgttgcacccgttgctggaggttttaaccc--ctcaagta	809
Db	577	CCAGTAGTCTCTAGCCCTTAACGATGATGACTACTACCCCTGTGCTATCGACCCCGTGCAG	518
Oy	810	acgaacctaaacgga tlaagtagacgcgccttggagactatctctgcagaag tgaacatcaaa	869
Db	517	CTGTAGCTAACGGCTTAAGTATCCCGCTGGGAGTACGTTCCCAAGATGAACACTCAA	458
Oy	870	ggaattgaacgggggtccgcacaagcggcttgagagatgltggtttaattcga tgaacccaa	929
Db	457	GGAAATTGACGGGGGCGCCGACAAAGCGGTGAGCATGTGTTTAATTCGATGCAAAAGCAA	398
Oy	930	aaacctaacctgggcttgacatgcatctgatac- atgtagagatatatgagcttcggs-	987
Db	397	GAACTTTTACAGGGCTTGACATGCGCGGAAATCCTTTGAAAGAGAGGGTGCCTTCGGGA	338
Oy	988	-cagattcaacagctgcctgcatagtgtgtgcctcaagctcgtgtcgtgagatgtttggttaagt	1046
Db	337	ACGGGAGACACAGGtGtGtGCATGGCTGTGtGCtACtGTGtGCCGTAAAGtTtTGGtTAAGT	278
Oy	1047	cccgcaacagcgcgaacccctatcgtatgttgccta- cctlaagttvggcactgtgaagaa	1105
Db	277	CCCCAACGAGACCGCAGMACCTCGTGTAGTTCAGTCATCGTGAATTTGGAAACCCCTGACAG	218
Oy	1106	actgcgggtgacaaacgcggaggaagcgggagatgaagatcaaatccctcatalgaccttaatgt	1165
Db	217	ACTGCGCGGTGTAAAGCCGAGAGAAAGGTGAAGATGACGTCAAGTCATCATGCCCCCTTATTC	158
Oy	1166	ccaggcgcaacacagctgcataaibgcga tacaagaggtgcgcacaactcgcacaagagagag	1225
Db	157	CTGTGGGAGACACAGCTGCTACAAATGGCGCGGAGCAAAAGGTGCGGATTCGCCGACGGGTGAG	98
Oy	1226	ctaattcttaaaagtcggtcccaagttcgaatttggggtcttgcacaactcgaccccatgaagtct	1285
Db	97	CTAACCCCAAAAACCCGCTCTCAAGTTGCGATTCAGAGCTGCACACTGCTCGCATGAAGCC	38
Oy	1286	ggaatcgctagtaatcgcgatcgaac	1311
Db	37	GGAAATCGCTAGTAAATCGCCGCTCAGC	12

RESULT	15
AQ969911/c	
LOCUS	
DEFINITION	AQ969911 676 bp DNA GSS 28-JAN-2000 LEURJ2R8Tf LERG Arabidopsis thaliana genomic clone LEURJ28, DNA sequence.
ACCESSION	AQ969911
VERSION	AQ969911.1 GI:6797612
KEYWORDS	GSS.
SOURCE	thale cress.
ORGANISM	Arabidopsis thaliana
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
	Kosidae; eurosoids II; Brassicales; Brassicaceae; Arabidopsis
REFERENCE	1 (bases 1 to 676)
AUTHORS	Buell,C.R., Lin,X., Pal,G., Barnstead,M., Bowman,C., Ullrichbach,T. Feldblum,T., Liang,F., Greasy,T. and Frisier,C.M.
TITLE	Genomic survey sequencing of landsberg erecta ecotype of Arabidopsis thaliana and identification of sequence-based polymorphisms
JOURNAL	Unpublished (2000)
COMMENT	Contact: Xiaoying Lin

```

source
1..676
/organism="Arabidopsis thaliana"
/strain="Landsberg erecta"
/db_xref="taxon:3702"
/clone="LERJ28"
/clone_id="LERG"
/notice="Organ: Leaf; Vector: pUC19uk; Total genomic DNA was sheared to 0.4-0.7 kbp before ligation."

BASE COUNT      133 a      207 c      177 g      159 t
ORIGIN

Query Match      24.3%; Score 359.6; DB 13; Length 676;
Best Local Similarity 76.7%; Pred. No. 2.3e-88;
Matches 517; Conservative 0; Mismatches 149; Indels 8; Gaps 6;

OY 755 gtaatcaacgacctaaagttgtctaccagttgtgtggggttllaacc--tcagtaag 812
Db 673 GTATGTCCTTAAGCCGTAAACCATGATGACTAGGCGCTGTGCTATGCAACCCCTGATGCTG 614
OY 813 aacctcaagatlaagtagaacccgcttgggagactatgtctgcacaagatgaactcaaga 872
Db 613 TAGCTTAAGCGCCTTAAGTATCCCGCTGGGAGATAGCTTGCAAGATGAACTCAAGAAG 554
OY 873 attgacggggttcgcacaaacggttggagcatgtgtttaaattcgaatgaaccacaaa 932
Db 553 ATTGACGGGGGCCCCACAAACGGGTGAGACATAGTGGTTTAATTCGATGCAAGCGAAGAA 494
OY 933 cctcaacctggagttacacatgatctcgaatc--atgtagagatatgaagccttcgg--ca 989
Db 493 CCTTACCAAGGCTTATGACATGCGCGCAATCCTTGAAGAAGAGGGGTGCTTCGGGAAG 434
OY 990 gattacaagggtgcgaatgtgtgtcgtcaagctcgtgtcgtgtagatgttggtaagtc 1049
Db 433 CGGAACACAGTGTGCATAGGCTGTGCTGACGTGCGCCCTAGAGGTGGTTAAGTCCC 374
OY 1050 gaagaagagcaaaccccatcgtatgtgtcacc--ttaagttggcactgtagaact 1108
Db 373 GCAAGAGGCGCAACCCCTCTGTTTAAATGTCACACCGTTGAATTTGAAACCTGACAGACT 314
OY 1109 gccggttgacaaacccggagaaagcggggaatgaacgtcaaaatccctctatgcttatgtca 1168
Db 313 GCCGCTGATTAACCCGGAGAGAGTAGAGTAGACGTCAATCATCATGACCCCTTATGCCCT 254
OY 1169 gggcacaccacgctgttacaatgcccgaatatcaaaaggttcgcaaccctcgaaaggagact 1228
Db 253 GGGCGACACACGTCTACATATGCCGGGCAAAAGGTTCGATCCGCGAGAGGTGAGCTA 194
OY 1229 atctctaaagtcggttcccaagtcgcatgtggaggtctgaacactcgaaacatgaatcga 1288
Db 193 ACTCCAAAACCCGTCCTCATGTTGCGATTGCAAGGCTGCAATCCGCTCGATGAAGCCGGA 134
OY 1288 atcgcctagtaatcgcggatcag--catgcccgggtgaatacgttcccgaaaccttgaacaa 1347
Db 133 ATCCGCTAGTAATTCGCGGTCACGCATACCGCGGTGATTCGTTCCCGGCGCTTGTACACA 74
OY 1348 ccggccgttacaacacacactgaatgcggggaagacccgaagtgtcttgttlaacgtlaagag 1407
Db 73 CCGCCCGTACACTMTGAGAGCTGGCCATGCCCGAAGTGT--TACCTTAACGCAAGAGAG 15
OY 1408 acagaactactaagg 1421
Db 14 GGAATGCCGAAG 1

```

FEATURES

Search completed: December 15, 2001, 02:33:34
Job time: 4972 sec

